

Two one-handed tilting-based writing techniques on a smartphone

Jari Laitinen

University of Tampere

School of Information Sciences

Computer Science

Supervisors: Erno Mäkinen and Poika Isokoski

April 2014

University of Tampere

School of Information Sciences

Computer Science

Jari Laitinen: Two one-handed tilting-based writing techniques on a smartphone

M.Sc. thesis, 85 pages, 14 appendix pages

April 2014

Text entry is a vital part of operating a mobile device, and is often done using a virtual keyboard such as QWERTY. Text entry using the virtual keyboard often faces difficulties, as the size of a single button is small and intangible, which can lead to high error rates and low text entry speed.

This thesis reports a user experiment of two novel tilting-based text entry techniques – with and without button press for key selection. The experiment focused on two main issues: 1) the performance of the tilting-based methods in comparison to the current commonly used reference method, the virtual QWERTY keyboard; and 2) evaluation of subjective satisfaction of the novel methods. The experiment was conducted using TEMA software running on an Android smartphone with a relatively small screen size. All writing was done with one hand only.

The participants were able to comprehend and learn to use the new methods without any major problems. The development of text entry skill with the new methods was clear, as the mean text entry rates improved by 63-80 percent. The reference method QWERTY remained fastest of the three throughout the experiment. The tilting-based technique with key press for selection had the lowest total error rate at the end of the experiment, closely followed by QWERTY. Interview and questionnaire results showed that in some cases the tilting-based method was the preferred method of the three.

Many of the shortcomings of tilt-based methods found during the experiment can be addressed in further development, and these methods are likely to prove competitive on devices with very small displays. Tilting has a potential as part of other interaction techniques besides text entry, and could be used to increase bandwidth between the device and the user without significantly increasing the cognitive load.

Key words and terms: M.Sc. thesis, smartphone, touch screen, writing technique, accelerometer, tilting, usability, user test.

Contents

1. Introduction.....	1
2. Evaluating text entry methods.....	5
2.1. Text entry rate.....	6
2.2. Error rate.....	7
2.3. Learning curves.....	9
2.4. Characteristics of text entry methods.....	9
3. Commonly used text entry techniques for mobile phones.....	11
3.1. Multi-tap.....	11
3.2. T9 by Tegic.....	12
3.3. Physical fullsize-keyboards.....	13
3.4. Virtual keyboards on touch screen.....	15
3.5. Alternative text entry methods on touch screen devices.....	16
3.6. Text entry rates with mobile text entry techniques.....	19
4. Tilting.....	20
4.1. Tilting as an input method.....	20
4.2. Design of tilting keyboards.....	26
5. Experiment.....	32
5.1. Goals.....	32
5.2. Participants.....	33
5.3. Apparatus.....	35
5.4. Design.....	39
5.5. Procedure.....	42
5.6. Pilot test.....	44
6. Results.....	45
6.1. Text entry speeds.....	45
6.2. Accuracy.....	46
6.3. Tilt efficiency rates with the Tilt keyboard.....	47
6.4. Subjective satisfaction of participants.....	48
7. Discussion and conclusions.....	54
7.1. Text entry speed development.....	54
7.2. Error rate variance.....	56
7.3. Problems regarding tilting.....	57
7.4. Questionnaire behaviour.....	57
7.5. Comments by the participants.....	58
7.6. Additional remarks about the results.....	60
7.7. Conclusions and future work.....	61
References.....	63
Appendices	

1. Introduction

The first text producing device for consumer usage was the typewriter which was invented as early as 1714. At that time there was little need to produce more text than could be achieved with just paper and ink, thus interest in the devices was slow to develop. The situation changed in the late 19th century with the beginning of industrialization. The need to produce even more text has grown steadily since [Silfverberg, 2007]. The first commercially successful typewriter was produced by E. Remington and Sons in the 1874 and it featured a QWERTY keyboard that was almost identical to the one used today [Yamada, 1980]. There were also many other types of typewriters available, like the index typewriters and typewriters with double keyboards, for lowercase and uppercase characters. Eventually, however, as touch-typing became more familiar with its superior speed, competitors started to lose the race. [Silfverberg, 2007]

There were attempts to improve the keyboard layout of the typewriters, as the original QWERTY layout had not been designed with touch-typing in mind. Perhaps the most familiar and most extensively studied of these is the Dvorak simplified keyboard designed by August Dvorak and William L. Dealey [1936].

The typewriters eventually evolved into electronic typewriters, which by the 1980s offered many features. When personal computers became successful around the same time, they simply adapted the keyboard from typewriters and these have not been changed much since [Silfverberg, 2007]. However, new text entry methods for personal computers have emerged, which are not based on the traditional QWERTY keyboard. For example commercial speech-to-text systems have been available since 1990s [Trewin and Arnott, 2007]. Also, different kinds of eye tracking text entry systems are available and have been studied widely [Majaranta and R  ih  , 2007].

There have also been attempts to improve text entry by reducing the number of keys on a keyboard by allowing several keys to be pressed simultaneously in order to produce different characters. These keyboards are called chording keyboards, and most familiar of these is perhaps the stenotype machine which is still used by court reporters in some countries. However, the most used text entry method on personal computers today is undoubtedly the QWERTY keyboard and it is unlikely that any other technique will replace it in the near future.

At the same time in the 1980s when personal computers made their way into the homes of ordinary people, mobile phones became commercially available. The first devices were used much like landline phones. However, as we now know, much has changed. Today, mobile phones are multi-purpose hand-held computers with little resemblance to the relatively simple voice communication gadgets that started it all. Currently, new mobile phones have features like email, video calls, photography, video recording, video on demand, GPS navigation, high-speed data access with web browsing capabilities, etc. The phones with more features and functions are often called smartphones to differentiate them from simpler and often less expensive models.

The term originates from 1997 when Ericsson described its concept device GS 88 "Penelope" as a "smartphone" [Stockholmsmartphone.org, 2010]. Although the device was never released to the public, it featured similar PDA-like features as the Nokia 9000 Communicator, which was released a year earlier [Nokia Press, 1996].

The first commercially available device that can be called a smartphone was the IBM Simon (see Figure 1). It was first demonstrated in 1992 and released to the public two years later, in August 1994. The phone was able to send and receive faxes, had a touch screen display and included applications common in PDA devices such as a calendar, a calculator, a notepad and more [IBM, 1994; BloombergBusinessweek, 2012].

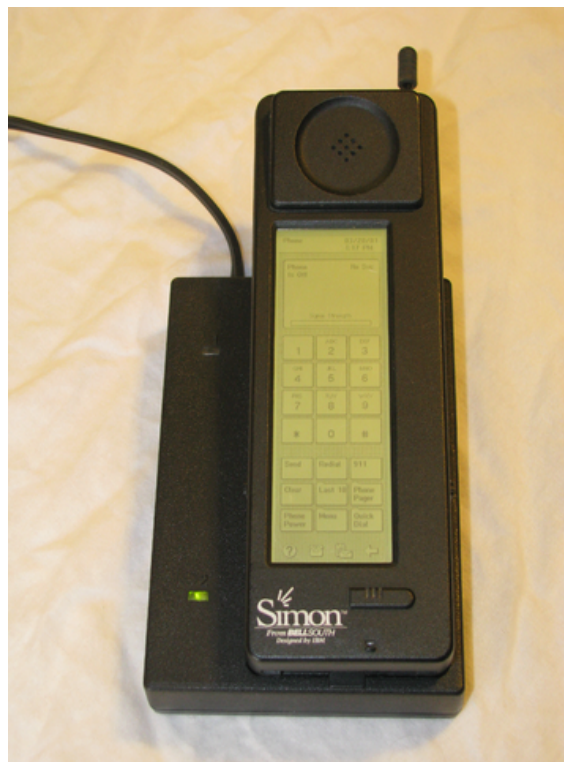


Figure 1: IBM Simon Personal Communicator. ¹

Even though the features mentioned above can be seen to be part of the unofficial definition of today's smartphone, perhaps the most important feature is the ability to run third-party applications. Application programming interfaces (or API's) are part of any operating system that runs today's smartphones. The ability to run third party applications is also mentioned on some definitions of the smartphone, like in PCMag's encyclopedia [PCMag, 2013].

From the perspective of this thesis the ability to run third party software makes it possible to create new text entry techniques for smartphones that utilize the touch screen and on-board sensors in different ways.

¹ Retrieved from http://commons.wikimedia.org/wiki/File:IBM_Simon_in_charging_station.png

The number of smartphones has increased rapidly [Gartner, 2009; Gartner 2013] and the International Data Corporation recently reported that more smartphones than ordinary mobile phones were sold for the first time in the first quarter of 2013 [IDC, 2013]. Mobile devices are not limited to phones, dumb or smart, as other categories such as netbooks, tablet PCs and personal digital assistants (PDAs) exist. However, sometimes it is hard to draw the line between devices like these as the tablet PCs have similar features as the smartphones while many of the new smartphones are equipped with the screen sizes close to smaller tablet PCs. For example Samsung recently released a new smartphone with a screen size of 6.3 inches [Samsung mobile press, 2013], while at the same time many smaller tables feature the possibility to make phone calls. PDAs on the other hand are often considered obsolete as the current smartphones include all their features while adding more.

As these devices have started offering a greater number of features, their popularity for business and entertainment use has increased. The International Telecommunication Union reported that there are over 6.8 billion mobile subscriptions globally as of February 2013 and that the number of mobile subscriptions is likely to exceed the world population in 2014. [ITU-T, 2013]

In the 1990s, mobile phone users started writing text messages to one another using the Short Message Service (SMS), which was gradually becoming more popular at the time. Today, mobile devices are used to write multimedia messages, answer emails, interact in social media and to do web searches – just to name a few. The one thing that is in common with all of these tasks is the need to enter text.

Text entry is one of the key features needed for any mobile device and many methods for it exist. These include ways to produce text using the traditional 12-key keypad, which was common on older devices, as well as other types of physical keyboards, like the full QWERTY keyboard. Newer touch screen based devices offer many different types of virtual keyboard layouts.

Text entry methods have been studied extensively from different perspectives with different types of input devices and varying target groups. A good text entry method is not easily defined. Most of the time, the literature makes references to speed, efficiency, ease of learning and ease of use in the text entry task.

Even though many competitive options exist, the QWERTY keyboard remains the most commonly used text entry method for Latin script in the mobile environment. However, small screen sizes and the touch screen often limits the text entry rate (see Table 1). To aid text entry on mobile devices, other solutions are needed. One solution is to add other modalities to the text entry task, which do not increase the cognitive load of the user significantly. Possible ways to do this is to use sensors that are already on-board with most of the current smartphones, like the accelerometer.

This thesis reports a study on the performance and usability of two new tilting-based text entry methods on mobile devices. Methods used to evaluate text entry techniques in the field of HCI are introduced in Chapter 2. Chapter 3 introduces some commonly used text entry techniques on mobile devices, including the virtual QWERTY keyboard used as a reference method in the experiment. Chapter 4 explains the use of tilting as an interaction modality in text entry and other uses and

introduces the original idea of the tilting keyboards and the early versions of their design. The experiment, apparatus, procedures and pilot testing are explained in Chapter 5. Chapter 6 presents the results of the experiment. Discussion and conclusions in Chapter 7 includes ideas for future work and possible new utilization possibilities for tilting.

2. Evaluating text entry methods

To evaluate the speed, efficiency, error rate or general goodness of any text entry method, we first need to know how these metrics are defined in a generally accepted way. Even though the prototype of a new text entry method is often tested by the developers, more formal testing is needed for reliable results. A generally accepted way in the field of experimental psychology is to follow the guidelines of the ROT test [Martin, 2008]. ROT refers to *repeatable*, *observable* and *testable*. "Repeatable" in the testing means that the report must describe the apparatus and the procedure used in detail, so that the test can be repeated by others having the report as their only guideline. "Observable" refers measurable responses, which in the case of text entry evaluations can mean the time to enter a phrase or the amount of errors. "Testable" refers to questions about the results; is the tested method faster than the other one?

In the field of text entry evaluation, experiments are usually conducted using randomly chosen participants. Participants write using the tested methods and often with reference method as well. The reference method is commonly known and has been tested in previous experiments. This way, the results of the new technique can be easily compared. The experiment usually takes place over several sessions, in which all the methods are used. This is needed as the text entry rate with the new technique is often quite low in the first sessions. Therefore, the results of the first few sessions are usually not comparable with the reference method, if the participants have prior experience using it. Multiple sessions also make it possible to evaluate how easy the new technique is to learn.

Even though the most natural way to test new text entry techniques would be to allow users to write freely, this would make comparison and evaluation harder. This is because when a user is allowed to write freely, the need to think "what to write" might affect the text entry performance, namely the speed. Secondly, identifying errors is hard because it is difficult to know exactly what the user intended to write. The third problem with free writing is the loss of control over the distribution of letters and words entered. Thus the entered phrases are not representative number of characters in the languages [MacKenzie, 2007; Wobbrock, 2007].

Text copying is usually preferred for evaluating text entry methods for these reasons [MacKenzie and Soukoreff, 2002]. MacKenzie and Soukoreff [2003] have developed a phrase set for this purpose. The 500 phrases in this set represent the letter frequency of English language and are generally short (mean length of 28.6 characters) so that they are easily memorable by the participants.

The research methods used can differ even within the text copying task. One of the issues is how and if the text correction is handled. In some studies, participants have been required to maintain synchronicity with the source text [Venolia and Neiberg, 1994; Isokoski and Käki, 2002; Evreinova *et al.*, 2004; Ingmarsson *et al.*, 2004] while in others, all the error correction mechanisms have been disabled [Matias *et al.*, 1996; MacKenzie and Zhang, 1999].

The method used in this thesis is the most widely used *unconstrained text entry evaluation paradigm* [Soukoreff & MacKenzie, 2001, 2003; Wobbrock & Myers, 2006]. In this methodology participants are able to enter any character and use backspace for error correction. A *presented string* is given to participant who enters it using the method under investigation – and the end result is called the *transcribed string*. No error beeps or other possibly intrusive effects are used. Participants are instructed to enter the string "as quickly and accurately as possible" and are free to choose what corrections are made. Log files are written which contain all the users input, including the corrections, all the keystrokes and other method depended input events. After the experiment, the log is parsed and interpreted to different measures. Measures used in the later chapters are explained in the following sections. [Wobbrock, 2007]

2.1. Text entry rate

Text entry rate is perhaps the most widely used measurement in text entry evaluations. Competitive text entry rate is often seen to be the key feature of any good text entry technique. Text entry rate can and has been measured in many different ways. The formulas used to calculate different measures often use the following definitions. The transcribed string that the participant has produced is denoted as T. T may contain letters, numbers, punctuation, spaces and other characters. It does not contain backspaces, however, and thus does not capture what has happened during the text entry process [Wobbrock, 2007]. From T we can get |T|, which is the length of the string T. IS is an *input stream*, which is another term used to describe the transcribed string. It contains all the inputs of the participant, including the backspaces. |IS| is the number of those inputs. The S is the time in seconds that the participant needed to complete T. It is measured from the first character to the last character.

The most fundamental metric is *characters per second* (CPS). In CPS, the number of characters (including spaces) in the transcribed string is divided by the amount of time it took to write the phrase. The formula for CPS is [Wobbrock, 2007]:

$$CPS = \frac{|T| - 1}{S} \quad (2.1)$$

Since the timing starts with the entry of the first character, -1 is subtracted from the length of the transcribed string.

One of the most widely used metric to report text entry performance is *words per minute* (WPM) [Wobbrock, 2007]. In WPM the length of the word is considered to be 5 characters, including spaces. This originates from as early as 1905 [Yamada, 1980] and has bases in the average word length in the English language, plus one for the space. WPM is computed using the length of

the transcribed string and the time it took produce it. The Formula for WPM is much the same as calculating the CPS. It is CPS multiplied by $60 \times \frac{1}{5}$ [Wobbrock, 2007]:

$$WPS = \frac{|T| - 1}{S} \times 60 \times \frac{1}{5} \quad (2.2)$$

Actual word sizes have been used by some researchers. However, this is not considered to be good practice as the length of the words affect the results and makes it hard to do comparisons between languages [MacKenzie, 2007].

The WPM measure does not take into account what the transcribed string contains compared to the presented string. To answer this problem *Adjusted words per minute* (AdjWPM) can be used. In AdjWPM the amount of errors remaining in the transcribed string affects the value of AdjWPM, and the actual value of AdjWPM depends on how much we want to penalize any amount of errors. [Matias *et al.*, 1996; Wobbrock *et al.*, 2006]

The CPS and WPM metrics do not take into account the amount of erroneous input during the transcription. The user may need to retype certain parts of the presented string many times before accepting it. To measure the efficiency of text entry activity from the user to the computer, another measurement can be used. The *keystrokes per second* (KSPS) characterizes the "data rate" of the text entry technique [Wobbrock, 2007]. The formula for keystrokes per second is:

$$KSPS = \frac{|IS| - 1}{S} \quad (2.3)$$

Gestures per second (GPS) is a metric similar to keystrokes per second. In GPS every action that contributes to the input stream is calculated, and then the GPS rate is calculated in the same way as KSPS. Gestures in this measure can mean any action like a keystroke, a tap with a stylus or tilting the phone when using tilt based text entry methods. It should be noted that even for the KSPS measurement, the "keystrokes" do not necessarily mean keystrokes, as they can be defined to be any

2.2. Error rate

Keystrokes per character (KSPC) is a measurement which can be used both as a performance and also as a characteristic measure of any text input technique. As the performance measure it is often used as a part of measuring errors in the input stream. KSPC is the ratio of the number of entered characters, $|IS|$ (including error corrections), and the length of transcribed string, $|T|$ [Soukoreff and MacKenzie, 2001; Wobbrock, 2007]. The formula for this is:

$$KSPC = \frac{|IS|}{|T|} \quad (2.4)$$

The KSPC performance measure does not make a distinction between backspaced characters that were correct and those that were not. For example, a participant might need to remove characters that are correct to add or change missing character in an earlier part of the phrase [Soukoreff and MacKenzie, 2004].

Minimum string distance (MSD) is a measure that can be used to compute the number of errors in the transcribed string. MSD gives the minimum "distance" between the presented string and the transcribed string by calculating the lowest amount of error corrections needed to get a perfect match between the two [Soukoreff and MacKenzie, 2001]. Correction operations to be used are "inserting a character", "deleting a character" and "substituting a character".

Because the accuracy during entry (KSPC) and the accuracy of the results (MSD) are not easily combined into one measure [Wobbrock, 2007], a unified error metric was developed by Soukoreff and Mackenzie in 2003. This unified error metric classifies all entered characters into one of four categories [Soukoreff and MacKenzie, 2003] These classes are:

Correct (C)	All correct characters in the transcribed text.
Incorrect-not-Fixed (INF)	All incorrect characters in the transcribed text.
Incorrect-fixed (IF)	All characters backspaced during entry.
Fix (F)	All backspaces.

With these four classes, three different error rates can be calculated. These error rates and the formulas used to calculate them are:

$$\text{Corrected error rate} = \frac{IF}{C + INF + IF} \quad (2.5)$$

$$\text{Uncorrected error rate} = \frac{INF}{C + INF + IF} \quad (2.6)$$

$$\text{Total error rate} = \frac{IF + INF}{C + INF + IF} \quad (2.7)$$

This unified error metric was also used on the experiment in this thesis.

2.3. Learning curves

To estimate the possible potential of any individual text entry method in terms of WPM after the initial testing has been done by the participants, the power law of learning model is often used [Card *et al.*, 1986]. The model is formulated as follows:

$$WPM = aX^b \quad (2.8)$$

Variable "X" presents "time" or the number of sessions in the experiment, "a" is the initial performance of the participant or group in question and "b" illustrates the steepness of the curve. With this model, it is possible to predict the performance of future sessions to a certain degree, especially if the coefficient of determination (R^2), which is used to indicate how well the data points fit to the statistical model, has a high value [Wobbrock, 2007].

2.4. Characteristics of text entry methods

Text entry methods can be characterized in many different ways. The keystrokes per character performance measure was explained in the previous section. Keystrokes per character can also be used as a characteristic measurement of any text entry method. In these characteristic measurements, KSPC describes the inherent efficiency of the text input method in question by measuring how many keystrokes or gestures are needed to generate a single character [MacKenzie, 2002]. However, even though a low KSPC or GPC might suggest a possible high text entry rate, it does not guarantee it, and should not be used alone to evaluate the efficiency of an input method in terms of speed, for example.

One key feature of any text entry method is the number of keys available for the text entry. This number is referred to as the *t factor* and is counted as the number of keys used to enter characters, plus one additional key for space. In English, the use of a space constitutes about 18% of all text entry [MacKenzie and Tanaka-Ishii, 2007]. For example, the traditional 12-key keypad on mobile phones has a *t factor* of 9, 8 keys for characters and one for space, and the English QWERTY keyboard has *t factor* of 27.

When the *t factor* is lower than the number of different characters to be entered, the text entry technique needs a way to handle the ambiguity of what character is intended. In a non-predictive method the user typically presses same key multiple times or uses some other pre-defined method from which to choose the desired character. In predictive methods, the system automatically handles the disambiguation and presents a list of ordered candidates from which the user chooses the intended character or word [MacKenzie, 2007]. Different kinds of models are used to sort suitable

character/word candidates based on what has been entered earlier by the user or in the language in general.

In text entry, the term "predictive" is used in two ways; a) to predict forward as word or phrase completion and b) to predict the user's intention [Tanaka-Ishii, 2007]. In text entry methods where disambiguation is needed, the term "predictive" is used in a narrow sense to actually predict the user's intentions. The term is used more in the narrow sense by companies and laymen, although text entry systems that are predictive in a broader sense also exist. In this thesis, the term is mostly used in a narrow sense if not otherwise mentioned. A further explanation of these models however is beyond the scope of this thesis.

3. Commonly used text entry techniques for mobile phones

In this chapter, some of the most common text entry techniques for mobile phones are introduced. All the text entry methods that are introduced here are used for the Latin script, although some are and can be used for other scripts as well. The methods introduced here are or have been at one point among the most used in mobile text entry. At the end of the chapter, text entry rates of the introduced text entry methods are shown in Table 1.

3.1. Multi-tap

Multi-tap is a text entry method which was widely used for mobile text entry when most of the mobile phones featured a 12-key numeric keypad. Even though the idea of the multi-tap method is quite old, it received new attention in the 1970s when the phone keypad for text entry was first considered [Smith and Goodwin, 1971].

The 12-key numeric keypad uses a key arrangement which has been standardized by the ITU Telecommunication Standardization Sector (ITU-T) [ITU-T, 2001] (see Figure 2). The standard ITU-T E.161 keypad consists of 12 keys, from which keys 2 to 9 contain characters, key 1 contains special characters and 0 is typically used to enter the space. Also, there are two additional keys, # and *.



Figure 2: 12-key phone keypad with the ITU-T E.161 basic Latin letter assignment. ²

In the multi-tap method, users enter a character by pressing the same key multiple times and cycling through the letters assigned to that key. For example, by pressing the key "7" three times,

² Retrieved from <http://en.wikipedia.org/wiki/File:Telephone-keypad.svg>

the letter "R" is selected. To enter the key, users simply pause for a while to enter the key or, alternatively, presses a different key. A special time-out key is available in many keyboards to speed up text entry in situations where there is need to enter multiple characters from the same key.

The multi-tap is a non-predictive method which has a keystrokes per character value of 2.0342 and *t factor* of 9 for the English lower case characters [MacKenzie, 2002]. Multi-tap was widely used to enter text messages when most of the mobile devices used the 12-key keypad. It is still available on most smartphones that utilize touch screens.

3.2. T9 by Tegic

The disambiguating phone keypad technique T9 was originally designed by Glover *et al.* [1995] is a predictive text entry method that resolves key sequences into their most likely words without requiring multiple key presses of the individual keys for individual characters. The technique is commonly used as an alternative for the multi-tap method and thus uses the same 12-key phone keypad layout, although the technique is not bound to any specific layout. This dictionary-based disambiguation method currently supports over 80 languages [Nuance, 2013].

When typing using the T9 technique, the user presses the key corresponding to a desired letter, and at the end of the word sequence is matched against device's dictionary. For instance, to enter the word "dog", a user would press the corresponding key series "364" on a device with an English dictionary installed. If multiple words match the sequence, the most likely candidate is shown and the user can browse through the alternatives. Different kinds of methods can be used to improve the most likely candidate selection, like the improved word list ordering by Gong *et al.* [2008] which uses context through semantic relatedness and a part-of-speech language model to improve the order of candidate words. However, even with the best of methods dictionary-based disambiguation methods can not predict words that are not in the device's dictionary, like abbreviations and slang words or when the user intends to enter numerals. Also, the user needs to keep focus on the display of the device constantly while writing to accept or reject the proposed words.

T9 by Tegic is not the only dictionary-based disambiguation method used in mobile phones, others include systems like eZiText by Zi and iTAP by Motorola, but it is the most widely used of such methods [MacKenzie, 2007] and has been installed on over 4 billion devices according Nuance [2013].

According to an analysis made by MacKenzie [2002] for English text, T9 has a keystrokes per character value of only 1.0072. However, testing only included words found in the device's dictionary, and as the value is so close to 1, the need to select another word than the first candidate was rarely used.

3.3. Physical fullsize-keyboards

The 12-key phone keypad is not the only physical keypad available for mobile devices that is used for text entry. Common alternatives are keyboards which mimic the layout of the QWERTY keyboards. The QWERTY layout was developed "well before 1887" for the usage of old typewriters [Yamada, 1980]. The design principle was not based on the efficiency of writing, but to minimize the mechanical jamming of early typewriters [Yamada, 1980]. Even though there are layouts that are designed with efficiency as the main factor, mainly the Dvorak simplified keyboard, which was mentioned earlier in chapter 1.1. [Dvorak and Dealey, 1936], the QWERTY layout has kept its popularity and was adopted as a standard by the American National Standards Institute in 1971 [Silfverberg, 2007].

One of the first commercially successful mobile devices that featured a physical QWERTY keyboard was the Nokia 9000 Communicator (see Figure 3, left), which was mentioned earlier. The device had a clamshell design, and when opened, the full keyboard and a dedicated display of 640x200 pixels were available to the user.



Figure 3: Nokia communicator 9000 and its successor, Nokia E7-00, released 2011. Both featured physical QWERTY keyboards. ³

³ Retrieved from [http://en.wikipedia.org/wiki/File:As_Time_Goes_By_\(Nokia_9000_Communicator_%26_E7\).jpg](http://en.wikipedia.org/wiki/File:As_Time_Goes_By_(Nokia_9000_Communicator_%26_E7).jpg)

The full physical QWERTY keyboard is not the only attempt to improve text entry rate when compared against the limitations of the 12 physical keys on the standard phone keypad. One example is the Fastap keypad for mobile devices by the former company Digit Wireless. It has the regular 12 number keys and characters A-Z are placed between these in alphabetical order (see Figure 4). This allows the phone to have all the keys available with a single press. Another is BlackBerry's reduced QWERTY keyboard (see Figure 4), which is a keyboard with the basic QWERTY layout. However, the number of keys is reduced, and two or more characters are placed on a same key, and to overcome the ambiguity, multi-tap or predictive techniques can be used [Silfverberg, 2007].



Figure 4: Left: LG 6190 ⁴ with the Fastap keyboard (left) and BlackBerry 7100v ⁵ with the Reduced QWERTY keyboard.

The physical keyboards on mobile phones can not compete in speed with the full size QWERTY keyboard on a desktop computer, as the typing usually happens with one or two fingers. Nonetheless, a study by Clarkson *et al.* [2005] showed that a relatively high text entry speed with a mini-QWERTY keyboard on mobile device can be achieved. Also, according to the study, such keyboards were seen as "marginally comfortable" by the participants. Still, physical QWERTY keyboards or any of their alternatives are rare these days, as most devices feature a touch screen and different kinds of virtual keyboards.

⁴ Retrieved from <http://www.esato.com/archive/t.php/t-76737,1.html>

⁵ Retrieved from <http://www.itreviews.com/research-in-motion-blackberry-7100v/>

3.4. Virtual keyboards on touch screen

Practically every new smartphone is equipped with a touch screen and has only few physical buttons. The few buttons that exist in these devices are used to control volume, the screen lock mechanism, the menu or to take pictures. For example, none of Apple's iPhones (see Figure 5) or Nokia's Lumia series phones are equipped with any sorts of physical keypad.

With devices like these, text entry is done using software based virtual keyboards, with fingers or with a stylus. Devices typically use the familiar QWERTY layout, although the virtual keyboard can also be, for example, a touch screen version of the 12-key keypad as well. This is one of the strengths of the virtual keyboards; as the keyboard is software based and uses the touch screen, they can be easily changed and customized, thus allowing users to easily switch between different kinds of region-specific keyboard layouts which have characters not included in the standard English keyboards. Different layouts are also possible and have been developed with efficiency in mind. These include layouts like Fitaly [Textware Solutions, 1999], OPTI [MacKenzie and Zhang, 1999] and ATOMIK [Zhai *et al.*, 2002]. According to studies, these layouts usually have higher text entry rates than the standard QWERTY layout on the same devices, but still have not been adopted widely for use, and QWERTY remains the default text entry layout on today's new devices. Because of the familiarity of the QWERTY keyboard, adopting new layouts takes a while for the user. In the studies where higher text entry rates with alternative layouts have been recorded, users have had time to learn the new layouts.

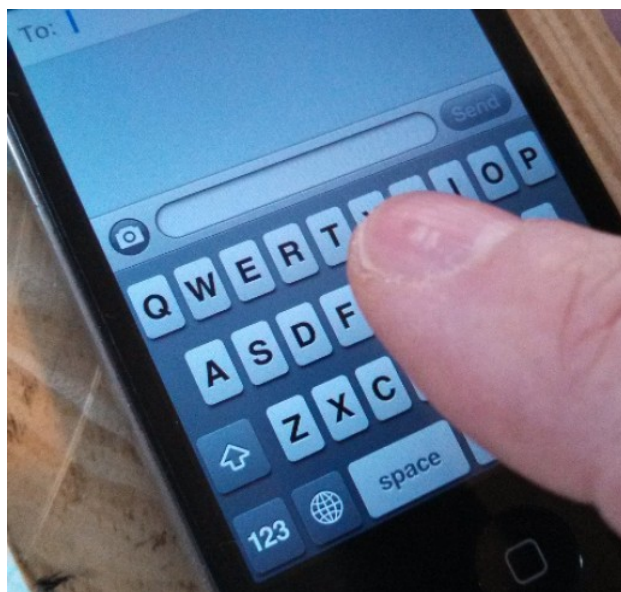


Figure 5: Virtual keyboard on Apple's iPhone.⁶

⁶ Retrieved from <http://dlrogers.hubpages.com/hub/Ending-the-QWERTY-Touchscreen-Tyranny>

Because the keys on virtual keyboards are intangible and can be quite small on some devices, different kinds of methods have been developed to aid text entry. Tactile feedback can be given when the user presses a button to indicate that the character has been entered. According to studies [Brewster *et al.*, 2007; Hoggan *et al.*, 2008] tactile feedback significantly improves finger-based text entry on touch screen based devices when compared to text entry without the feedback. Pakkanen *et al.*, [2010] have studied the usefulness of different kinds of tactile feedback on mobile devices. According to their studies haptic stimuli can be used to deliver more complex information to the user than just the button press confirmation and could be therefore be used to aid text entry in the future. Another widely used aid for text entry is to show an enlarged version of the currently selected character above the finger. This, together with the ability to slide a finger over the keyboard and selecting a character only when the finger is released, instead of pressing the key, can make the character selection process easier. As the feature requires constant visual focus for the user, it has its possible drawbacks as well. Martin *et al.*, [2009] studied the benefits of the over-headed character and found that there was no clear difference in terms of text entry or error rate with or without the over-headed character. However, user responses indicated that the feature was appreciated by the participants. Also, different kinds of auto-correction mechanisms have been developed to aid text entry and are often utilized with the use of virtual keyboards.

3.5. Alternative text entry methods on touch screen devices

The touch screen provides other alternatives for text entry besides tapping the keys with a finger or a stylus. Handwriting recognition is a relatively old technique and was already available on devices like the Newton Message Pad, which was released in 1993 [Apple-History, 1996]. Handwriting recognition gives a familiar way for text entry and can happen on different levels as the recognition process can consist of the recognition of individual characters to words and even sentences [Silfverberg, 2007]. The technique, however, has several problems as well. Because of the similarity of different numbers and characters, and multiple ways users can produce them, hand writing recognition systems often have a hard time interpreting what was intended. To overcome this problem, new character sets have been developed. In these character sets, individual characters can correspond to symbols, for example, which are well separated from each other, thus helping machine recognition. The disadvantage of techniques like these is the fact that the user must remember the new set of characters, which often do not resemble their Roman counterparts [Tappert and Cha, 2007]. One solution to make the recognition easier is to limit the way characters are drawn by a predefined order of the strokes, similar to what is used to write Chinese characters. Another way is to make each character different enough from one another. If the way characters are drawn is kept close to their original Roman counterparts, users do not need to learn a totally new character set and the learning phase of the new techniques is relatively short. One example of a technique like this is the EdgeWrite alphabet by Wobbrock [2005].

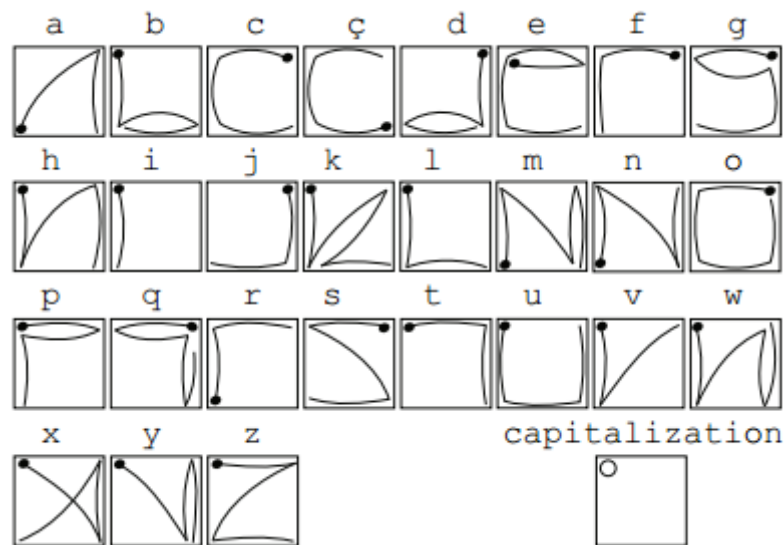


Figure 6: Basic alphabet set of the EdgeWrite.⁷

The EdgeWrite is an unistroke text entry method, which has a character set that closely mimics the Roman characters (see Figure 6). This makes characters easier to learn and memorize while maintaining the easy recognition levels in the system.

Handwriting recognition techniques are rarely used today on smartphones, although some devices do support them. This is likely due to the goodness of the virtual keyboards which often offer lower error rates and similar or faster text entry rates, according to studies [Zhai and Kristensson, 2003; Költringer and Grechenig, 2004; Luo and John, 2005].

Another alternative to utilizing the possibilities of touch screens is to provide a handwriting like text entry technique combined with the familiarities of a graphical keyboard layout. In these techniques, a user slides their finger or stylus over the keyboard layout crossing all the characters included in the word, starting from first and ending with the last character. After the user lifts the finger from the keyboard, the system resolved what the intended word was. In techniques like these, tapping the keys is usually also possible. One of the most studied of these systems is ShapeWriter by Zhai and Kristensson [2006]. ShapeWriter supports the regular QWERTY keyboard layout as well as the optimized ATOMIK layout mentioned earlier.

⁷ Retrieved from <http://depts.washington.edu/ewrite/downloads/EwChart.pdf>

Many similar text entry techniques exist. Perhaps the most famous of them is the Swype keyboard (see Figure 7), which was originally developed by Swype Inc. In this technique, a user slides their finger or stylus through every letter of the intended word, starting from the first and ending with the last, lifting only between words. Like in ShapeWriter, the system uses error-correction algorithms and language model to resolve the intended word. Swype supports a number of localized QWERTY layouts and languages, including Finnish. It is available on Google Play and is pre-installed on some Android devices.

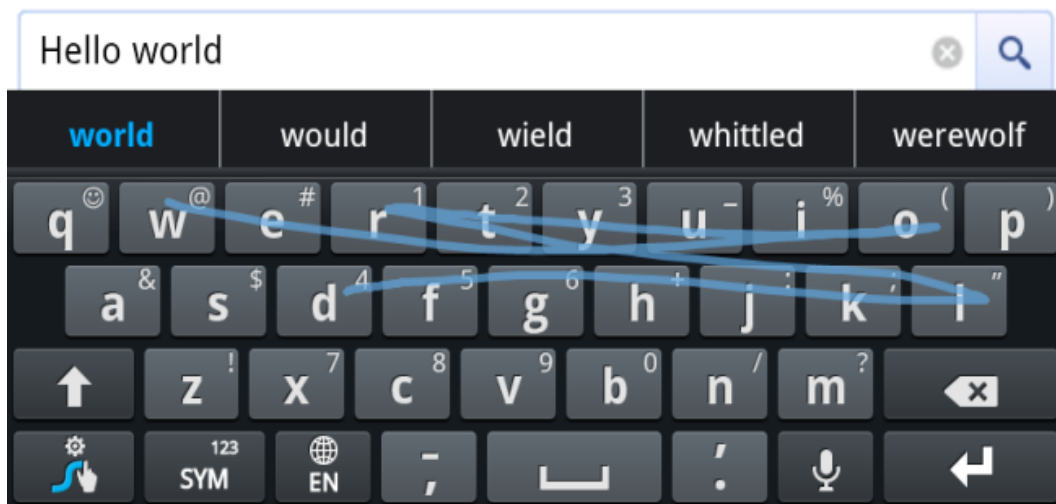


Figure 7: Swype running on Android. The word "world" has been written and the system presents possible alternatives.⁸

⁸ <http://www.howtogeek.com/106643/how-to-type-faster-with-the-swype-keyboard-for-android/>

3.6. Text entry rates with mobile text entry techniques

Many factors affect the text entry rate of any particular text entry technique in any study or experiment. These may include the number of participants and their age, gender and prior expertise with the tested methods or devices. Also, the choice of the devices influences the results, as well as the settings used in the experiment. Thus the text entry rate for any text entry technique depends on many different factors. The results in Table 1 are taken from different studies and are not directly comparable to one another, but are presented for illustrative purposes only. The text entry rates mentioned here are all taken from studies where their performance was tested on text-copy writing task in the English language.

Text entry technique	Text entry rate (WPM)	Source(s)
EdgeWrite	5.9 – 6.6	[Wobbrock et al., 2003]
Fasttap	7.1 – 8.5	[Cockburn and Siresena, 2003]
Multi-tap	7.2 – 15.5	[James and Reischel, 2001; MacKenzie <i>et al.</i> , 2001; Butts and Cockburn, 2002; Wigdor and Balakrishnan, 2003; Silfverberg, 2007]
T9	9.1 – 20.4	[James and Reischel, 2001; Silfverberg, 2007]
Virtual QWERTY	13.6 – 18.5	[Költringer and Grechenig, 2004; MacKenzie <i>et al.</i> , 2009]
Swype	58	Guinness world record 2010
Physical mini-QWERTY	31 – 61.7	[Clarkson <i>et al.</i> , 2005]

Table 1: Text entry speed estimates with different text entry methods on mobile devices.

4. Tilting

This chapter focuses on tilting as an input method. The first section describes the working principles of tilting as an input method and briefly introduce its practical applications. The applications introduced are mostly text entry related, but also include other utilizations of tilting as part of another input method. The second section presents the idea and the design principles of the tilting-based novel text entry techniques.

4.1. Tilting as an input method

New mobile devices, namely smartphones and tablets, are equipped with a wide range of sensors. One of these sensors, which can be found in almost any new mobile device, is the *accelerometer*. Most accelerometers are *micro electro mechanical sensors*, or MEMS, which use a small mass etched into the silicon surface of the integrated circuit and are suspended by a small beams. As acceleration is applied, a force develops which then displaces the mass. These beams act as a spring, and air is often used as a damper. The amount and direction of the force is transformed into an electronic impulse. The accelerometer's sensor is typically capable of sensing velocity and position, tilting, and orientation on one to three dimensions. [Elwenspoek and Wiegerink, 2001]

The accelerometer is often used to sense the orientation of the phone, which makes it possible to change the screen layout between horizontal and vertical automatically. The accelerometer's sensor has also been utilized in some text entry techniques by way of adding to the bandwidth between the user and the system. In this chapter, some of these text entry techniques that use tilting as part of the text entry process, are presented. Also, some other practical implications are presented.

4.1.1. Tilting as an input method on arm-worn devices

Tilting has been evaluated as an input type on arm-worn based devices (e.g. wrist watches). One of the examples of these is TiltType, designed by Partridge *et al.* [2002]. It uses four physical buttons combined with eight compass directions for tilting. Thus, the trade-off between the number of buttons and tilting directions favours few buttons and many tilting directions.

TiltType was tested and developed for wrist watch type devices, although it can be used on other devices as well. To enter a character, the user presses one of the four physical buttons on the side of the device and the display then shows characters corresponding to that button. While holding the button, the user tilts the device to select desired character from one of eight tilting directions and then releases the button to confirm the selection. To select the ninth character, located in the middle,

the user does not tilt the device, but instead just presses and releases the button. See Figure 8 for a complete list of character mappings.

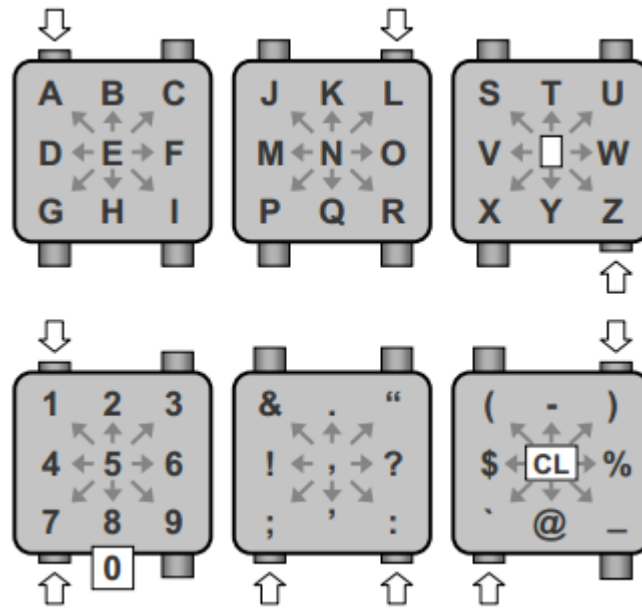


Figure 8: Mapping of characters in TiltType [Partridge *et al.* 2002].

The device requires two hands to operate because of the position of the physical keys, although operating it with one hand is possible when the device is strapped to the wrist. However, this requires tilting the whole forearm, which makes the technique difficult to use according to its authors. Controlled studies were not conducted by Partridge *et al.* [2002], instead the prototype device was informally tested by 50 people who were able to write their own name or enter a few characters using the technique.

Tilting has also been used as part of an interaction method on other arm-worn devices. Daniel Fallman [Fallman, 2002] designed a PDA-based arm-worn device which uses tilting as the means to scroll the screen. A screen bigger than the display area is navigated by tilting the device. Although the technique is not for text entry per se, it demonstrates the possibilities of tilting as an input method when the screen size is limited.

4.1.2. Tilting on mobile phones

As mentioned earlier in section 4.1., the accelerometer can be used to sense the orientation of the device. It has therefore been utilized as part of the screen orientation switch mechanism on numerous devices. Gaming is an area where the accelerometer has frequently been utilized. Device tilt for example can be used to control the speed or direction of a moving object. Tilting has been utilized as an input method for other applications as well, like menu selection. Because of the low

cost and many uses it offers, the accelerometer has been an integrated part of most mobile phones for quite some time now.

The usefulness of tilting of the device has been experimented with as part of text entry techniques on mobile devices. TiltText is an example of a tilting-based text entry technique for mobile phones by Wigdor and Balakrishnan [2003]. TiltText is a technique designed to be used with the traditional 12-button phone keypad, which was commonly in use when the technique was developed. The technique uses a combination of button presses on the keypad and the tilting of the device to determine the desired letter, and it is not dependent of the language used. When compared against traditional multi-tap methods, which have three to four characters assigned to a single button, TiltText has a single character assigned to a single button-press-and-tilt -event, thus removing ambiguity. Characters are entered by pressing a key and simultaneously tilting the phone in one of four directions (left, forward, right or back). Compared to TiltType, the technique therefore has fewer tilting directions than buttons (see example in Figure 9 for available characters when button 7 is pressed).

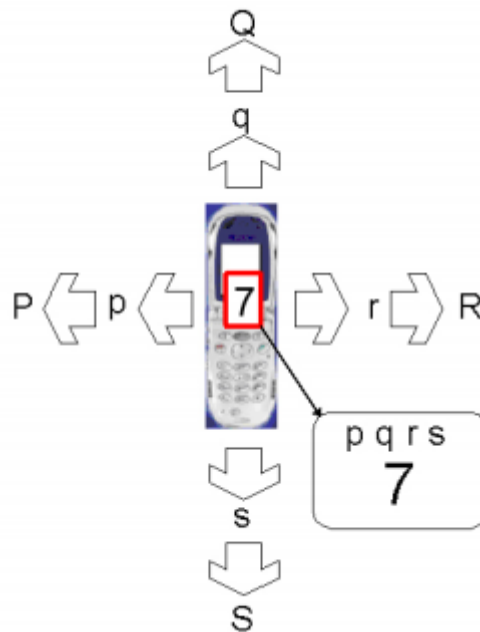


Figure 9: TiltText: When the user presses keypad button 7, eight different characters can be entered [Wigdor and Balakrishnan 2003].

TiltText uses the same buttons for characters as the traditional 12-button mobile phone keypad. To enter the first character of any key, the user presses the key and tilts the phone to the left. To enter the second character, the phone is tilted forward, and to enter the third character, it is tilted to the right. The only time the user needs to tilt back is when button 7 is pressed. It contains four different characters on the ITU-T E.161 standard keypad. Pressing any key without tilting enters the

number on the key. To enter an upper-case character, the magnitude of tilting can be used for disambiguation; when the phone is tilted more, the upper case character is entered.

Wigdor and Balakrishnan conducted a study where they tested their design against the multi-tap method with ten participants. According to the study, initially the text entry rates were almost identical, 7.42 WPM and 7.53 WPM for TiltText and Multi-Tap, respectively. At the end of the study, the rates were 11.04 WPM for Multi-tap and 13.57 WPM for the TiltText method. The error rate with the TiltText method was higher than with the multi-tap method – 11% for TiltText against 3% for Multi-tap. As the procedure in this experiment was to correct every error so that the end phrase is completely correct, the text entry rate still remained higher for TiltText regardless of the higher error rates.

The technique was further developed by Wang *et al.* [2006] as part of their studies to utilize the camera on mobile phones to detect motion in a technique called *TinyMotion*. Using TinyMotion they constructed a new version of TiltText which they called *Vision TiltText*. The technique is similar to the original TiltText, but uses the phone's camera to sense tilting and also detect movement to left, for example, has same effect as tilting to left.

Another extension that uses TiltText's original idea is the *Solving Collision and Out of Vocabulary Problems in Mobile Predictive Input with Motion Gesture*, or SHRIMP. It is a method developed by Wang *et al.*, [2010] which combines tilting and dictionary-based disambiguation. Dictionary-based disambiguation, like T9 which was presented earlier, typically has two problems. When it is used with the standard 12-button phone keypad, more than one word may correspond the same numeric sequence and users need to select the right one. Another is that out of vocabulary words (i.e. words not present in the built-in dictionary) are hard to enter. SHRIMP uses traditional dictionary-based disambiguation, but adds Vision TiltText features for any single character selection.

SHRIMP works just like T9 and the user presses a key which corresponds to a character. When the whole word is entered, the system shows which words correspond to the given numeric sequence. SHRIMP adds an option to tilt the phone when entering any character, and the user can do this for all or none of the characters in the current word. Even when one or two individual characters are entered using the TiltText method, thus making that character unambiguous, the amount of possible corresponding words drop dramatically. Particularly, entering out of vocabulary words is easier as the user can use the Vision TiltText method to tilt every character. The experiment by Wang *et al.*, [2010] showed that the SHRIMP method was significantly faster than the Multi-Tap or the Vision TiltText methods alone, and showed an average speed of 12.1 WPM. Also, the method was generally liked by the participants.

Tilting is a supported modality for many input methods. One of the most studied and best known is the *Dasher method*, developed by Ward *et al.* [2000]. Dasher (see Figure 10) is a text entry technique which supports multiple platforms and modalities, including a computer mouse, a Wii remote and an eye gaze, just to name a few.

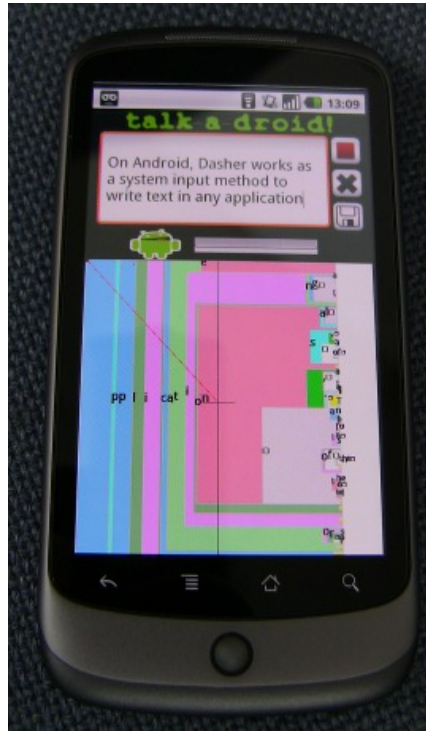


Figure 10: Dasher running on an Android device.⁹

In Dasher, the user chooses one letter at a time from the screen while the system uses a probabilistic predictive model to anticipate the next likely character based on previous choices. Next possible letters and letter combinations are located on the right side of the screen and the user selects the desired one by way of zooming into it. The zoomed character crosses the center of the screen and thus becomes the selected character.

Currently, Dasher is available for most mobile operating systems, including iOS and Android. It supports tilting as the means to select a characters. Text entry rate with Dasher using tilt as the modality has not been studied. However, according to Ward and MacKay [2002], the text entry rate using Dasher on a computer with a mouse showed text entry rates up to 34 WPM for expert users. Up to 25 WPM was measured with an eye tracker as the input device after just an hour of practice.

Text entry is not all about producing character after character. Many times, there is a need to edit or format the text later on. Dearman *et al.*, [2010] have studied four different modalities in terms of text selection and text entry. The studied modalities were touch, tilt, speech and foot tap. The results showed that tilt was the fastest method to select a text target, but also produced the highest amount of errors when compared to touch – which had the highest overall text throughput of

⁹ Retrieved from <http://www.inference.phy.cam.ac.uk/dasher/websource/english/MobileDasher.html>

all the modalities tested. Also, the tilt and foot tap methods were not liked by the participants as coordinating the selection efficiently was felt as awkward by some of the participants. Touch and speech were commented as the most natural modalities in the study.

Tilting has also been studied as part of an input method on mobile devices for other usages than just text entry. Turunen *et al.* [2009] developed and tested a mobile phone as a means to control a home media center. The system accepted multimodal inputs, and among them was the ability to use tilting to move the selection on the screen up and down. As tilting was part of the gesture based navigation used on the system, its usefulness alone in this context is hard to evaluate. However, this study shows yet another possibility of using a tilting as part of an input method.

Medryk and MacKenzie [2013] compared user performance and enjoyment during gameplay between touch and tilt input methods. In their experiment, they used a game called Bit.Trip Beat for this purpose. The game is reminiscent of Pong, where the user manoeuvres a paddle vertically forward and backward. The goal of the game is to bounce back the incoming pixels.

All 18 participants played the game with both input methods. The participants only played the first level of the game. The dependent variables were level completion time, error rate and final score of the level. The touch input resulted in better scores for all the dependent variables, and the differences were statistically significant. The improvement with practise was greater for tilt-based input (6.2% for tilt input and 1.1% for touch input). The post-experiment questionnaire shows that 72% of the participants preferred the tilt-based input over the touch input. The cited main reason was that "it was more challenging and more engaging to use the accelerometer and tilt the device to control the paddle and bounce back pixels".

MacKenzie and Teather [2012] have also studied tilting as an input method. Instead of applying tilting as an input method for some specific task, the study focused more on fundamental questions about the use of tilting as an input method in general. Their main interest was human performance, e.g. how well users can control an on-screen object using tilting as an input method, and what parameters of tilt influence the performance on different settings. They also tried to confirm whether the tilt input follows Fitts' law, as input from devices such as mice and joysticks, and what the throughput (i.e. an Index of Performance in bits/time in Fitts' law) for tilt-based interaction is.

In their study, they used the ISO 9231-9 standard multi-directional position-select task for non-keyboard input devices. In the task, the users select targets which are arranged in a layout circle. The position of the target changes according to a predictable pattern. Three different target amplitudes were used in the experiment (125, 250 and 500 pixels). Tilting was used to move the cursor (a ball of 20 pixels in diameter), and there were three different sizes of target circles (40, 60 and 100 pixels in diameter). They used two different settings to select the circle; "first entry" where the circle is selected when the ball first enters the destination circle, and "dwell time" (500ms), where user needed to keep the cursor inside the destination circle for it to be selected. Also, four different tilt gains were used (25, 50, 100 and 100). Tilt gain refers to the sensitivity of the tilting. The experiment had 16 participants, and it was performed using a Samsung Galaxy Tab 10.1. A

short survey followed the experiment where Likert-scale questions were used to assess physical and mental effort, as well as comfort and ease of use.

According to their results, a tilt gain setting between 50 and 100 is optimal for movement time and throughput. Also, the first-entry selection is faster than the dwell selection, and was also favoured by the participants. Maximum tilt angles ranged from about 2° to 13°, depending on condition. According to their study, tilt as an input method conforms to Fitts' law. However, the throughput is low – about 2.3 bits/s for first-entry selection and 1.2 bits/s for dwell selection.

If the tilting alone would be used as an input method on a text entry technique, dwell selection would likely to be the only way to select a character. If used with the keyboard layouts that are designed to be efficient (see Section 3.4.), estimated text entry rate would be relatively low. If we assume the average pointing distance between two keys on an optimized layout to be two keys, we can get $\log_2(A/W+1)=\log_2(2/1+1)=1.58$ bits per character. For dwell time of 500ms this would mean around 9 WPM $((1.2/1.58)*60/5)$. If the first entry could be used, which would be same as dwell time of 0 ms, theoretical text entry rate would be 17.5 WPM $((2.3/1.58)*60/5)$.

If tilting would instead be used to lower the Index of Difficulty by allowing targets to be bigger and the selection task would be done by touch input, which was shown to outperform tilting [Medryk and MacKenzie, 2013], higher text entry rates could be achieved.

4.2. Design of tilting keyboards

In this chapter, I will present the idea and the design principles of the tilting-based novel text entry techniques. Working principles of both novel techniques are shown, as well as the early prototype and ideas behind the techniques.

4.2.1. Early design and usability issues

The touch screen has become the standard of mobile device input methods. In general, very few new phones or tablet devices exist that use a physical keyboard. Even though the screen size of today's smartphones is on the rise, many users still prefer using devices with smaller screen sizes. Also, smartwatches like Pebble, by Pebble technology, SmartWatch 2, by Sony and Galaxy Gear, by Samsung, which have all been released commercially in the past year, feature small touch screens.

The small screen makes text entry using the virtual QWERTY on-screen keyboard hard or even impossible. The reason for this is the size each button, which can be very small. Also, as the buttons on a touch screen are intangible, they do not offer the same kind of haptic feedback to users as the physical buttons do. Text entry on large touch screen devices is not without its problems, either. If the users need to enter text using one hand only, the size of the device can make reaching some of the keys difficult. Situations like these may lead to higher error rates and lower text entry rates.

Theoretical text entry speeds of button pressing based text entry techniques, on touch screen or otherwise, can be estimated using a model that utilizes Fitts' law for the aimed pointing portion of the interaction. [Fitts, 1954; MacKenzie, 1992].

To overcome some of the challenges that text entry faces on touch screen based phones, I needed a way to make hitting the buttons easier while still maintaining the unambiguous nature of the full QWERTY keyboard. The easiest solution for hitting the small buttons would be to make them bigger. However, this would soon lead to issues with the limited amount of screen space. The solution I used for this was to make keyboard layout change on the fly, thus allowing the use of bigger keys and a single character per key at any time. I chose tilting as the modality for changing the on-screen keyboard layout. I expected that tilting would not add to the users' cognitive load significantly, and that text entry would be possible without the need to constantly think about how to enter the next character.

The first version of the Tilt keyboard consisted of four pages of letters, each with 12 buttons. Also, keys for the space, the backspace, and some function keys were always available, see Figure 11.

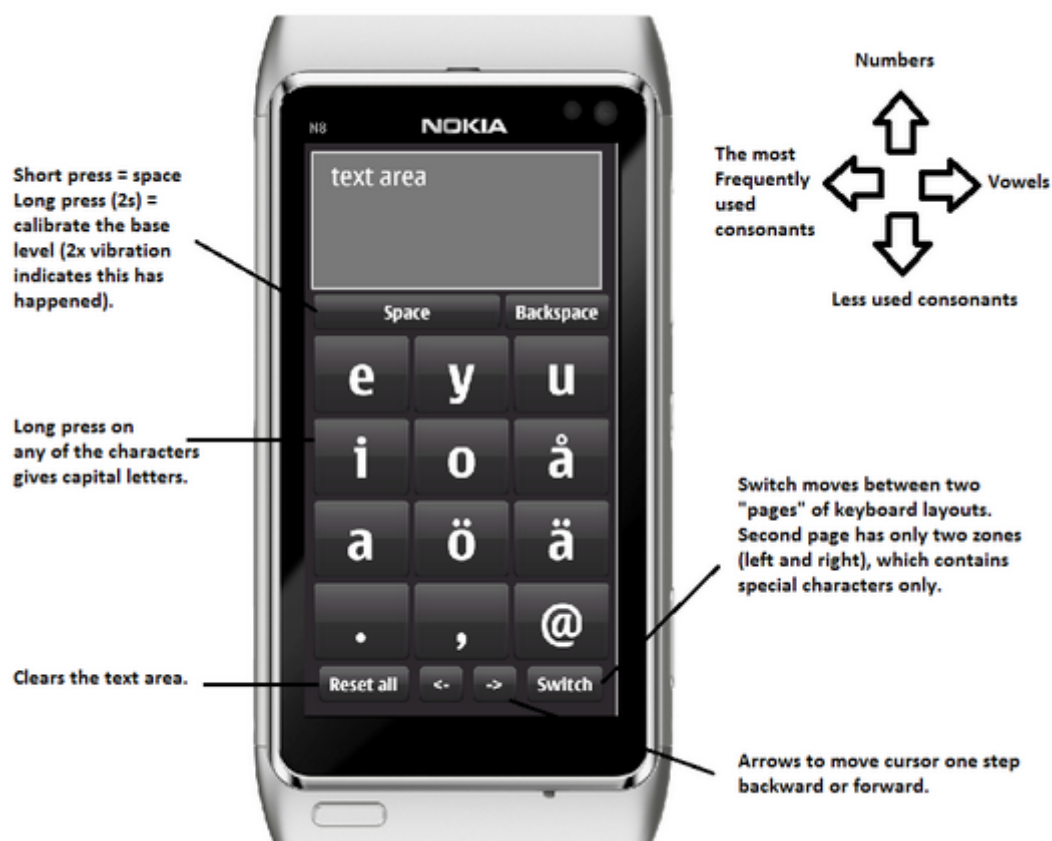


Figure 11: The first working prototype of the Tilt keyboard running on Nokia N8.

The layouts on keyboard pages on the first working prototype were set so that they were easily remembered by the users. All the vowels were located on the right page, all the numbers were on

the upper page, the most common consonants in the English language [Algoritmy.net, 2013; data-compression.com, 2013] were on the left page and less used ones on the down page. Between the layout pages was a central position, which we call the *neutral position* from now on. The neutral position is the position of the phone where the user feels it is most comfortable during writing. The neutral position works as a "safe zone" where writing can happen and the keyboard layout does not change. The neutral position is set automatically when the Tilt keyboard is loaded, but can be set manually by the user at any point. This is done by a long press of the space button.

When the user wants to change the layout, he/she tilts the device 20 degrees or more from the neutral position in one of the four tilting directions. 20 degrees was chosen so that users would not change the layout by accident while writing, but only when they had clearly tilted the device. The characters are entered by pressing the buttons on the touch screen. This can be done while the phone is tilted, or on or near the neutral position. The screen layout switching is always compared to the neutral position and the user must keep in mind where it is located.

After the initial testing of the prototype, it was decided that the number of keys per page would be reduced to nine, making it possible to increase the size of the keys for each page. Also, the extra function keys were removed, so that all the space available could be reserved for the changing layouts. The space and backspace keys remained unchanged, and are always available. The amount of keys remaining easily allowed all the characters in the English language to be set for three pages. The page layouts were optimized to minimize the need to tilt after each character. In the new layouts, the most common characters were located on the right page, the second most common on the left page, the third most common on the up page and punctuation marks and special characters on the down page. The characters inside the pages were placed based on the syllable frequency of the English language [Stefan Trost Media, 2013]. The letters that mostly go together are located close one another on the layout pages. According to a study by Rahman *et al.*, [2009], the pronation/supination axis (i.e. left and right) has a better resolution than the ulnar/radial axis (i.e. up and down) while performing tilting-based tasks. This supports the layout as the user mostly needs to tilt the phone on the pronation/supination axis.

It was also decided that it would be interesting to also explore a tilt-based text entry system for a devices without a touch screen. For this purpose another tilting-based text entry method was developed. This keyboard was named as Tilter. It uses the same keyboard layouts as the Tilt keyboard, but is more tilting-based. Tilter shares some similarities with TiltType, designed by Partridge *et al.* [2002], that was presented earlier in section 4.1.1. Tilter uses only two physical buttons instead of four as in TiltType. Detailed working principles of Tilter can be found in sections 4.2.4. and 4.2.5.

As both keyboards use the same layout, it was possible to evaluate the impact of tilting as a modality in text entry and as an input method in general. Also, the two different methods allowed for the evaluation of the usefulness of touch input as key selection when compared to techniques where selection happens with tilt input without the dwell time. The next two sections explain the working principles of both tilting-based text entry techniques in detail.

4.2.2. Tilt keyboard

The Tilt keyboard uses a combination of touch screen button presses and the tilting of the phone to change the keyboard layout. When the user initiates a text entry task and the Tilt keyboard starts, one of the four keyboard pages, the start page, is shown. The start page contains the most common characters used in the English language. The start page and the other pages are shown in Figure 12. 11 keys are shown on each page. One key is always the backspace and another is the space. The function of the other keys depends on tilting.

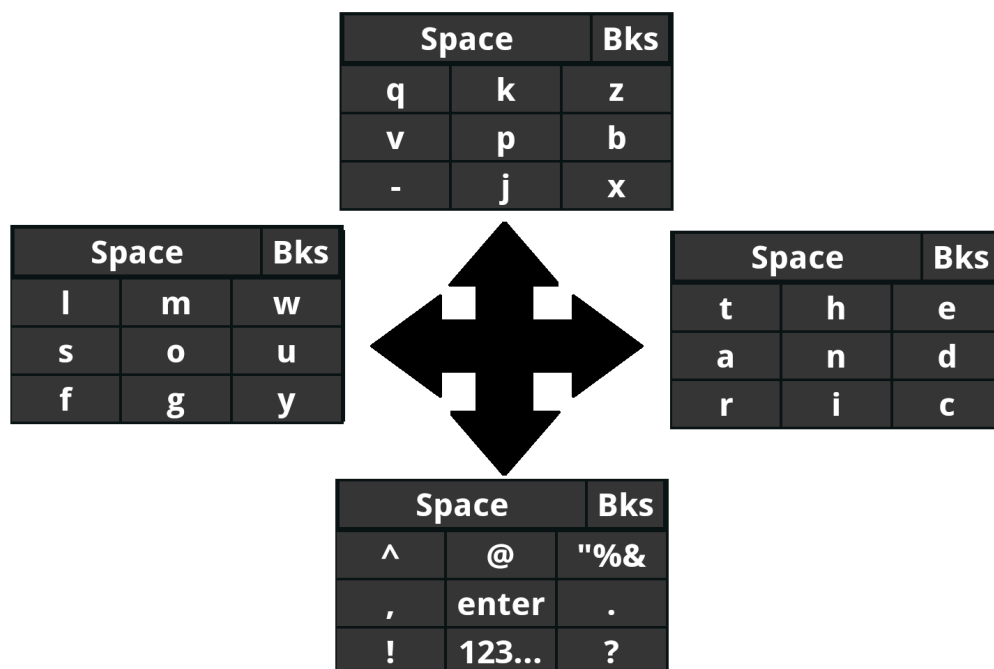


Figure 12. The Tilt keyboard layout and tilting directions.

4.2.3. Using the Tilt keyboard

When the phone is tilted, a keyboard page, each with different sets of characters/functions, is shown. Tilting directions are left, right, up and down. The most common characters can be found on the right page, the second most common on the left page and the least common ones on the up page. The down page contains special characters and function keys. A character or function is selected by pressing a button with a finger.

When the Tilt keyboard starts the orientation/position of the phone is stored automatically. This is called the neutral position. The neutral position is a "safe zone" between the keyboard layouts where the layout does not change. Tilting is compared against the neutral position so that when the phone is tilted in one of four tilting directions from the neutral position, the layout changes

accordingly. When typing, the phone should be tilted in the direction where the desired character can be found and then returned to the neutral position for selecting the character/function. Although characters can be entered while the phone is tilted, preliminary testing showed that the user might lose awareness of the neutral position if not returned to it while writing.

The neutral position can be reset at any time by pressing the volume up button at the left side of the phone while the phone is in the position that is intended to be the new neutral position. Resetting the neutral position should be done if the layout changes rapidly or not at all and thus is difficult to control.

4.2.4. Tilter

Tilter uses the tilting mechanism of the phone as the main input method to enter characters and function keys. The touch screen is not used at all, although the volume up button works as an activation key for the tilting mechanism while the volume down key works as the backspace key. When the user initiates a text entry task and Tilter starts, one of the five keyboard pages, the start page, is shown. The start page has the space key and eight empty buttons. All the pages are shown in the Figure 13.

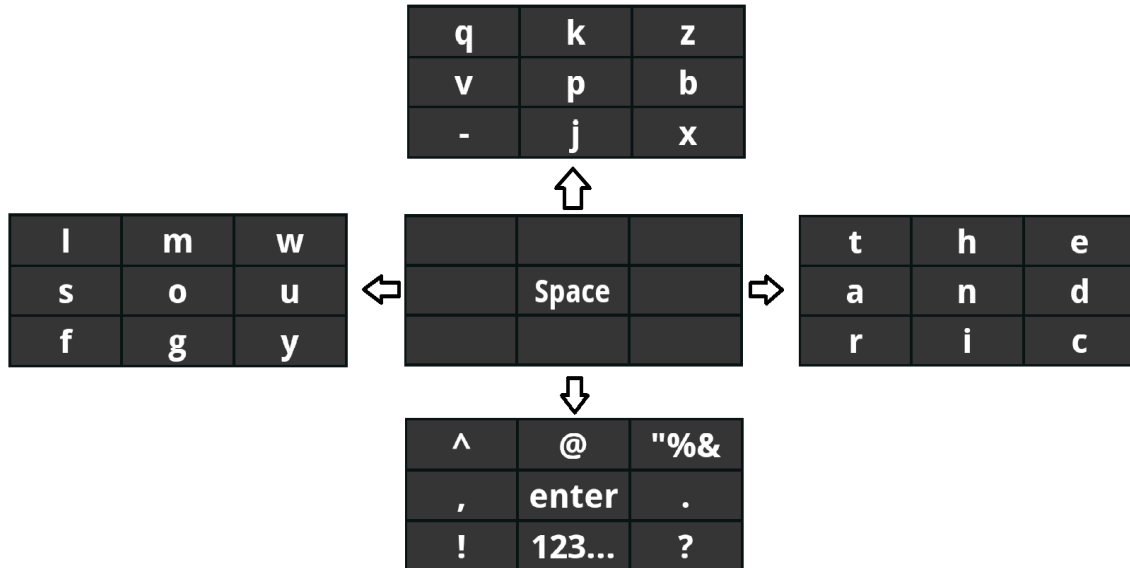


Figure 13. The Tilter keyboard layout and tilting directions.

4.2.5. Using the Tilter

Entering a character with Tilter is done in four phases: First, the volume up button on the side of the phone is pressed and held down. Second, the phone is tilted toward a desired keyboard page while still holding the volume button down. Third, the phone is tilted in direction of the desired character (see Figure 14 for an example of a highlighted character). Finally, fourth, the volume button is released to select the character and to return to the first page.

t	h	e
a	n	d
r	i	c

Figure 14. The character 'e' has been highlighted.

The volume up button is located on the left side of the phone that was used in the experiment. Tilting directions are left, right, up and down. The most common characters can be found on the right page, the second most common characters on the left page and the least common ones on the up page. The down page contains special characters and function keys. To input the space character, the volume button is pressed down and released without tilting.

One of the nine buttons shown must be selected to continue after selecting a page. The mechanism to recover from an erroneous page selection is to enter a character and then press the volume down button to erase the entered character. When the volume up button is released, Tilter shows the start page.

5. Experiment

In this chapter, I describe the experiment that evaluated the usability and likeability of the tilting-based novel keyboards. The goals of the experiment are described in the first section. The second part presents the participants. In the third section the apparatus used in the experiment is introduced. The fourth section explains design of the experiment. The fifth section explains the procedure and in the sixth section the pilot test setup and early results are presented.

5.1. Goals

The main goal of the experiment was to investigate the performance of the two novel tilting-based text entry methods against the currently common text entry methods in terms of text entry speeds and error rates. Also, the goal was to evaluate subjective satisfaction of the used methods at the early phase and at the end of the experiment. The behaviour of the users in terms of tilting was also analysed as part of the experiment.

As the tilting-based techniques are new and therefore none of the participants had prior knowledge of writing using such a technique, text entry speed was expected to be slow at the start. This is mainly due to visual search, as the layouts were not familiar to the users. One of the key factors affecting the steepness of the learning curve was the users' ability to use the tilting to change the layouts and familiarize themselves with them. As the text was written using one hand only, participants were not likely to have much experience using the reference method in this way, either. The one hand only use was chosen as the technique was developed to be used in less than ideal circumstances, like when the surrounding is moving (e.g. when on a bus) or when the user has only one hand available (e.g. while driving a car, carrying groceries or due to disability). Another situation when the technique is likely to outperform traditional touch based text entry methods is when the display of the device is very small. One hand use allowed for better evaluation of the usefulness of tilting in the text entry when compared against the reference method QWERTY keyboard in the same conditions.

Another interesting question was the users' ability to cope with the heavily tilting-based Tilter method. The technique is clearly different than the other methods used in the experiment and the approach is quite uncommon among text entry methods in general as well. Even though the technique was expected to be slower than the other techniques used in the experiment, it still has its potential use. While the other techniques rely on the use of a touch screen, the Tilter can be used with any device that has an accelerometer, a display and minimum of two physical buttons. The technique can be, in theory, used with extremely small display sizes as well. The only limiting factor will be the user's ability see the layouts. It was interesting to see how well the participants

were able to use this technique in the beginning and what were the differences in the text entry speed and total error rates at the end of the experiment with other tested methods..

The participants' evaluations of the methods was the third question. It was expected that the individual participant's performance with the methods might reflect the evaluation. As for the novel techniques, it was interesting to see how the participants would evaluate the techniques especially in terms of usability and likeability and whether the evaluations changed towards positive and negative during the experiment.

5.2. Participants

The experiment had 6 participants, 4 male and 2 female. Before the experiment, all participants filled a background questionnaire (Appendix 1 English and Appendix 2 Finnish). The questionnaire contained questions about the prior knowledge of different text entry methods and how the participants wrote on their own mobile devices (for more details see Table 1 and Appendix 1 and 2). Both language versions contained the same questions. One of them was given to each participant according to his or her native language.

All the participants took part in the experiment without knowledge of payment or compensation. In the end of the experiment, a single entry movie ticket was given to all participants to compensate for the time spent in the experiment.

	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Participant 6
Age	39	25	29	21	50	25
Gender	Male	Male	Female	Female	Male	Male
Handedness	Right	Left	Right	Right	Right	Right
Education	Master's degree / Ph.D.	Bachelor's degree	Master's degree / Ph.D.	Gymnasium / Grammar school / High school	Master's degree / Ph.D.	Master's degree / Ph.D.
English skill	Native speaker	Excellent	Good	Excellent	Excellent	Excellent
As technology user	2	1	1	2	1	1
Screen size of used device	1.8"	3.7"	3.7"	3.7"	4"	3.2"
Average daily use	3	3	3	2	5	3
Familiar methods	2/5	4/5	4/5	3/5	3/5	4/5
Default text entry method	Multitap	Virtual QWERTY	Virtual QWERTY	Virtual QWERTY	Virtual QWERTY	T9
Reasons why producing text is hard on a smartphone	2, 3, 4	4	2, 4, 5 (Auto-correction)	2, 4	2, 3	4

Table 2: Background information of the participants.¹⁰

The age of the participants ranged between 21 and 50 years and most were between 21-30 years. Five participants were right handed, one was left handed. The left handed participant preferred to use his right hand in the experiment.

The majority of the participants chose "Master's degree or Ph.D." as their education level. For the English language skill, one of the participants was a native English speaker while the most selected item was "Excellent". The participants were also asked about the size of their primary mobile device in the questionnaire. Most of the participants had a relatively common size screen on their mobile device, as the marked screen sizes were between 1.8 and 4 inches.

The participants were also asked to estimate how many words they write with their mobile device on an average day. The response alternatives were 1) "Over a thousand words per day", 2) "Hundreds of words per day", 3) "Clearly more than ten, but less than a hundred", 4) "Around ten

¹⁰ The number value on some of the questions indicates the chosen item in the multiple-selection e.g. 1 = the first available item.

words per day", 5) "A few words per day" and 6) "I do not write with my mobile device". Most of the participants answered "Clearly more than ten, but less than a hundred", while one answered "Hundreds of words per day" and one "A few words per day".

The participants were also asked about their previous experience levels with different mobile text entry methods as well as what method they currently primarily used for text entry on their mobile device. If the participants were not familiar with the methods based on names alone, they were instructed to ask more details about them. Also, a picture of the Multi-tap method was shown from the phone's screen when the participants were filling the question in hand.

Most participants used the virtual QWERTY keyboard for text entry on their own devices, while one used Multi-tap and one T9 as their main text entry methods. Some users were still occasionally using other methods than what they normally used. The knowledge of different methods varied between the users.

The participants were also asked if they felt that producing text using a smartphone is hard and if so, why. Multiple selections for the response alternatives were possible and all the participants marked at least one item. The two most popular answers were "keys are too small" and "avoiding typing errors is hard". None of the participants thought that screen size would be an issue as "screen is too small" was never selected. One participant mentioned "autocorrect" under the "other reason" field.

5.3. Apparatus

The apparatus in this experiment consisted of several different parts. This section contains detailed description of the hardware and software solutions used in the experiment. First, the chosen test device is presented, followed by the software used for the testing. The last part of this section explains the reference keyboard used.

5.3.1. Device

The TEMA application [Castellucci and MacKenzie, 2011] and the tilting keyboards software were running on a Samsung Galaxy Y (GT-S5360) (See Figure 15), which is an Android-based smartphone made by Samsung. The phone has a 3 inch (76.2mm) TFT LCD capacitive touch screen with a resolution of 240 * 320 pixels and a pixel density (PPI) of 133. The phone was running the Android OS version 2.3.6 and did not have any extra applications aside from TEMA and the tested text entry methods (IMEs).



Figure 15: Samsung Galaxy Y (GT-S5360). ¹¹

5.3.2. Text entry metrics for Android

Several different alternatives were considered as a means to measure the text entry performance. As the tilting keyboards record their own logs, by analysing those logs, text entry speed and error rates could have been measured. The problem with this approach would be the writing part itself. How to present test phrases to the participants and where the writing process would happen.

Ultimately, an independent program for just this purpose was chosen. Text entry metrics for Android (or TEMA for short) is an application to calculate text entry performance and accuracy on any portable device running an Android OS 1.5 or later. It has been used to evaluate the performance of text entry methods in previous studies. TEMA automatically stores all the input characters and events to two log files on the phone. The logs contain text entry speeds, error rates, the presented and transcribed characters and the elapsed time. The event log contains time-stamps of each input event on each test phrase. When running TEMA on the test devices, the use of IMEs is possible, which makes it easy to test the performance of novel techniques against the old ones.

¹¹ Retrieved from http://commons.wikimedia.org/wiki/File:Samsung_Galaxy_Y_S5360_run_Android_2.3.6.jpg

TEMA can also take into account possible interruptions during the testing, like when the user presses the menu button by accident and has to return back to software to continue writing. TEMA automatically deducts these pause times from the transcription time. [Castellucci and MacKenzie, 2011]

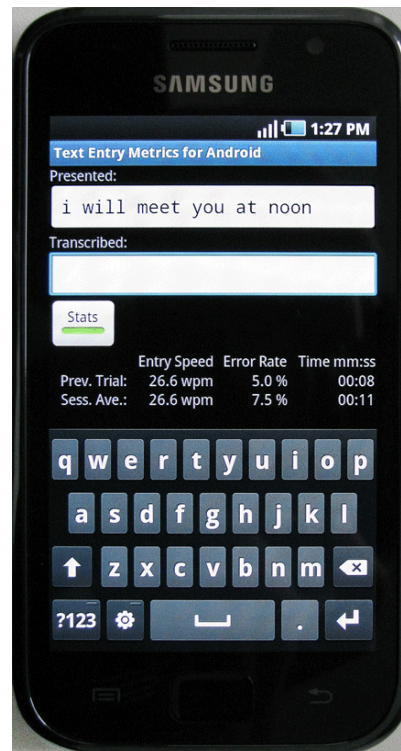


Figure 16: Text entry metrics for Android is an application for Android OS 1.5 or higher to calculate text entry performance and accuracy on any portable device [Castellucci and MacKenzie, 2011].

While testing is running, the phrase to be entered is located in a text field in the upper part of the screen and under it is an empty field where participants' input will be entered. Space for the IMEs is under these. The statistics part, which is visible in Figure 16, was not seen by the participants during testing. Measuring of text entry of each test phrase starts when the first character of a phrase is entered and ends when the user presses enter at the end of the phrase.

TEMA uses MacKenzie's and Soukoreff's 500 test phrase corpus. Phrases are moderate in length and easy to remember. They also represent the target language, which in this case is English. The length of the phrases ranges between 16 to 43 characters while the average phrase length is 28.61 characters [MacKenzie and Soukoreff, 2003]. Each test phrase is randomly chosen from the set. It was possible for the same phrase to appear multiple times when 10 random phrases were written with each method in every session. All the characters in the 500 phrase set are in lowercase and do not include any numbers or punctuation marks. Figure 17 shows five randomly chosen phrases from the set of 500 phrases from TEMA:

you are not a jedi yet
starlight and dewdrop
arguing with the boss is futile
the cotton is high
the most beautiful sunset

Figure 17: Five randomly chosen phrases from MacKenzie's and Soukoreff's 500 test phrase corpus [MacKenzie and Soukoreff, 2003].

5.3.3. The QWERTY keyboard

In this experiment the used QWERTY keyboard was the default keyboard of the phone (Samsung keyboard). Locale settings were set to "Suomi" (Finnish), thus the keyboard layout included three extra Nordic characters, "ä", "ö" and "å". Other than that the layout follows the traditional QWERTY look. Nordic version of the QWERTY keyboard was chosen mainly because most participants were native Finns. One participant was originally from New Zealand, but he had lived in Finland for many years and was also familiar with using the Nordic QWERTY keyboard. Figure 18 shows the QWERTY Nordic keyboard and the QWERTY UK keyboard layouts as they appear on the screen of the device used in the evaluation on a vertical layout.

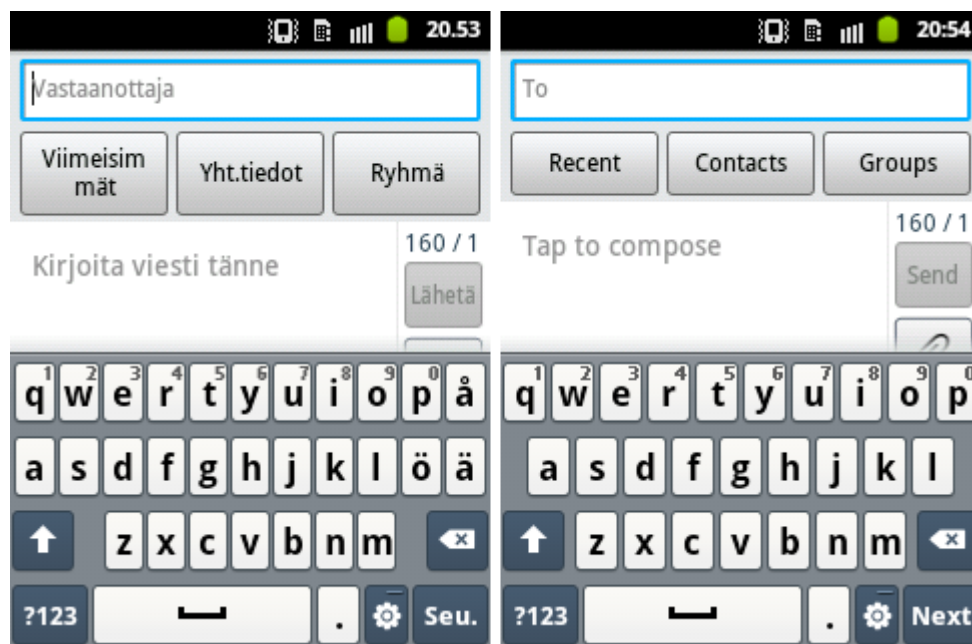


Figure 18: QWERTY Nordic (left) and QWERTY UK (right) on a Samsung Galaxy Y phone.

The phone's screen width was 240 pixels. In the QWERTY Nordic method each character key is 19 pixels wide and in the QWERTY UK method each character key is 20 pixels wide. This means that the character keys are 5 percent wider in the QWERTY UK version of the keyboards. There is also a difference between the spacing of the keys. In the Nordic layout there is one pixel of empty space between the keys. In the UK layout there are two pixels. It should be noted that even by pressing the seemingly empty space between characters can often result in a different character being entered than the one which was nearest of the touch point. This is due the integrated statistical language models embedded into the text entry systems of the current mobile devices. The main function of these systems is to predict or guess what is the most likely next character that the user intended to input.

The QWERTY keyboard was chosen as the reference method mainly due to its familiarity and also because it has obtained the status of being the default text entry method for almost every touch screen device. Older methods, like the Multi-tap, were also considered. However, comparison to the method that is currently more commonly used was considered to be more meaningful. For the same reason, methods which could produce higher text entry rates after enough practice, like the Swype method introduced earlier in chapter 3, was not chosen. Also, the method is not the primary choice for most users. From now on the Nordic QWERTY keyboard is simple referred as the QWERTY keyboard unless otherwise mentioned.

5.4. Design

The experiment was a 3 x 6 within subjects factorial design and the factors were:

- Text entry method (QWERTY, Tilt keyboard, Tilter)
- Session (6 sessions)

The dependent variables were entry speed, accuracy and efficiency of tilting. The writing order was balanced by Latin square. Table 3 below shows the orders used. E.g. participant 1 started with order 1 and ended with order 6 and participant 4 started with order 4 and ended with order 3. The column "Duration (minutes)" shows the completion times of the sessions.

Session number	Order	Duration (minutes)
Session 1	QWERTY, Tiltkb, Tilter	40-80min
Session 2	Tilter, QWERTY, Tiltkb	20-40min
Session 3	Tiltkb, Tilter, QWERTY	20-35min
Session 4	QWERTY, Tilter, Tiltkb	20-30min
Session 5	Tilter, Tiltkb, QWERTY	18-30min
Session 6	Tiltkb, QWERTY, Tilter	30-45min

Table 3: Sessions, order of the methods, and duration range of each session.

In each session, participants wrote ten phrases with all three text entry methods. Session 1 lasted around an hour, sessions 2-4 about 30 minutes and sixth and final session about 40 minutes – varying between participants. The sessions were scheduled so that there was always a minimum of two hours between sessions for each participant. The participants also had the opportunity to take breaks between methods and phrases.

The phrases were randomly chosen (with replacement) from a 500 phrase set included in TEMA. The collected data includes the values from the logs by TEMA (see 6.3.2.) as well as logs by the Tilt keyboard. From the Tilt keyboard logs, tilts over the neutral zone boundary and reset of neutral zone were calculated. From these values the values for the dependent variables were computed. Text entry speed was calculated by TEMA:

"...by dividing the length of the transcribed text by the entry time (in seconds), multiplying by sixty (seconds in a minute), and dividing by five (the accepted word length, including spaces). The entry speed was averaged over the ten phrases and reported in words-per-minute (wpm)." [Castellucci and MacKenzie, 2011]

Accuracy was also automatically calculated by TEMA. Accuracy is reported as the total error rate (TER) which characterizes general input accuracy and is the sum of the corrected error rate (CER) and the uncorrected error rate (UER). These are reported as a percentage of average of ten phrases written with each method in a session. [Castellucci and MacKenzie, 2011]

Tilting efficiency was calculated by dividing the actual number of tilt events by the minimum number of tilts needed to enter each presented phrase. Every tilt that is more than 20 degrees from the neutral position, and therefore changes the keyboard layout, is considered a "tilting event" (see 5.1. and 5.2.1. for more information). *Characters per tilt efficiency*, which does not take into account the pressing of the enter button, was calculated from the 10 phrases that were entered.

The QWERTY UK method was tested at the end of the last session for reference. The tested novel keyboards did not include the Nordic characters that were present in the QWERTY keyboard

used in the evaluation. The UK version of the QWERTY keyboard was tested to see if there is any major difference in text entry speed or accuracy between the two. Each participant wrote 10 phrases using it.

5.4.1. Questionnaire for subjective satisfaction

The NASA Task Load Index (later NASA-TLX) is a multi-dimensional scale which is designed to obtain workload estimates from users while they are performing the task at hand or immediately afterwards. It was developed by the Human Performance Group at NASA Ames Research Center during a three year research project which involved more than 40 laboratory simulations. Its original application was in aviation, but it has been used in many different kinds of studies since its original manual was released in 1986. It has been cited in over 550 studies and a search from Google Scholar for the term "NASA-TLX" reveals around 8000 different articles [Hart, 2006]. [Original NASA-TLX manual, 1986]

The original NASA-TLX uses a two-part evaluation procedure consisting of both sources of load (weights) and magnitude of load (ratings) and has 6 different factors which are used to evaluate the workload of given task. These factors are *Mental demand*, *Physical Demand*, *Temporal Demand*, *Performance*, *Effort* and *Frustration*. In the rating part of the NASA-TLX, users fill out a 21-point scale questionnaire consisting of each 6 factors. In the second part each subject compares these factors pairwise based on their perceived importance. The number of times each factor is chosen is added to its weighted score, which is later multiplied by the scale score from the rating part of the tool, and then divided by 15 to get a workload score which will be the overall task index of the performed task. [Original NASA-TLX manual, 1986]

Some studies have used the so called "Raw TLX", which does not include the pairwise comparison. Some researchers suggest that the shortened version might increase experimental validity over the full one [Bustamante and Spain, 2008].

This experiment used a questionnaire to evaluate subjective satisfaction of the text entry methods. The questionnaire was based on the 21-point rating part of the NASA-TLX. The questionnaire did not include the pairwise comparison of the original NASA-TLX. Some of the factors were modified to be better suited for the text entry evaluations. The NASA-TLX has been modified in many different studies in the past [Complete TLX Publication List, 2013]. The full list of the factors used in this work can be seen in Table 4 below and in the questionnaire from Appendix 3.

Also, in this experiment original paper-and-pencil -version of the questionnaire was used, although online and stand-alone-software version of the NASA-TLX that simplify the data gathering process do exist. The study by Noyes [2007] suggest that the use of the paper-and-pencil version might lead to less cognitive workload than filling a questionnaire on the computer screen, which was shown to effect the evaluation.

Factor	Description
Mental demand	How mentally demanding was the task?
Frustration	How insecure, discouraged, irritated, stressed, and annoyed were you?
Task success	How successful were you in accomplishing what you were asked to do?
Speed of use	How fast you feel you were using the text entry method in question?
Ease of use	How easy it was to use the method in question?
Learnability	How easy it was to know how to use the method in question?

Table 4: Factors and their descriptions in used questionnaire that was based on the NASA Task Load Index questionnaire (Appendix 3).

5.5. Procedure

In the first session the procedure and goals of the experiment were explained to the participants. Their only knowledge before this was that they would be taking part in an experiment where different kinds of text entry methods would be tested and participating would require to take part in six different sessions. The first session would last about an hour on average, sessions 2-5 around 20-40 minutes and last around 30-45 minutes. After explaining the procedure and the goals, the participants were asked to fill in the permission form (Appendix 4) and the background questionnaire (Appendix 1 and Appendix 2). As part of the instructions it was explained that in this experiment writing would happen with one hand only with all three methods. This was also reminded at the start of each session.

The experiment took place in a small and quiet room which was free of interruptions. Lights were dimmed so that they would not reflect from the screen of the device, which was noticed to have a possible effect in the pilot test phase. The phone was held in a portrait orientation and the automatic orientation change was set to off. All the sounds were set off and screen brightness to highest possible setting. The phone did not have network available during the testing – no SIM card present and Wi-Fi turned off. The phone did not have any cables plugged during testing or any screen covers. The phone was booted before each session and logs were transferred from the device after each session. The phone's screen was also wiped between each method with a microfibre cloth

so that the finger marks would not affect the performance. Also, the participants had the opportunity to use a tissue papers if needed to wipe sweat of their hands, for example.

Before the actual testing, the novel techniques were introduced to the participants. This was done by giving the participants written instructions on paper on how the technique worked (Appendix 5 and Appendix 6 in English and Appendix 7 and 8 in Finnish) just before the method was to be tested for the first time. Instructions were also given verbally and the participants had the opportunity to test writing with these methods for a few minutes before the actual testing began. Instruction papers, that showed the key layouts, were taken away before the actual testing. There were no specific instructions given for the QWERTY method, other than to use it with one hand only. The participants did not have opportunity to test the method before the testing began. The participants were free to hold the phone in the way they felt most comfortable. The use of fingers were not limited in any way. All the participants used the thumb to enter text.

All participants were instructed to type the test phrases as fast as possible while avoiding typing errors. The participants had the opportunity to take as long breaks as they wanted between methods or phrases, but if they had started writing a phrase, they were instructed to continue until it was completed. Also, as a part of the instructions in the earlier experiments with TEMA [Castellucci and MacKenzie, 2011], the participants were instructed to correct writing errors when they happened and not if they later noticed error on an earlier part of the current phrase. It was emphasized to the participants that the experiment did not test their performance, but the performance of the text entry methods.

Also in the first session the participants were asked to close or mute their own mobile phone to avoid interruptions while the testing took place. Some instructions were repeated during sessions, like the need to reset neutral position for the Tilt keyboard at before each test phrase, as was instructed on methods instruction paper (Appendix 5 and Appendix 6 in English and Appendix 7 and 8 in Finnish).

The modified NASA-TLX based questionnaire (Appendix 3) was filled right after each used method in sessions 1 and 6. Participants were instructed to mark at the top of the lines how they felt. Also, at the end of sessions 1 and 6 short interview were conducted, which consisted three main questions: "Which method you liked best", "Which felt fastest" and "What would you use" and possible other question based on participants performance or earlier comments. Participants also had the opportunity to comment freely between test phrases and methods.

The QWERTY UK method was tested in the last session, after the participants had used all other methods and answered all questionnaires and interview questions. After QWERTY UK was used, participants were asked to comment freely how the method felt compared to the QWERTY method used in the experiment. All the participants completed all six sessions.

5.6. Pilot test

A short informal pilot testing was conducted with two participants who did not take part in the actual experiment. One of the goals of the pilot test was to test the functionality of the novel methods, e.g. to discover possible bugs that were not found during the development and self-testing. The apparatus used in the experiment was also tested. Also, the level of instructions needed for the use of the tilting-based method were tested. The goal was not to test the performance of the methods. The participants had the opportunity to give comments freely and questions were asked to clarify how well the participants understood the given instructions regarding the use of the methods and testing procedures. This information was later used to fine-tune the instructions and test procedures, and the decision to add the NASA-TLX based questionnaire was made based on the findings of the pilot test.

The participants filled the same background questionnaire that was later used in the actual tests, which was also tested in this way. The content of some of the questions was slightly altered for the actual experiment, like adding one more option for the question about daily text entry amounts of participants.

The results of the pilot test in terms of text entry speed and accuracy were consistent with the results of the actual experiment that was done later in more controlled conditions. No mentionable problems with the novel text entry methods software, procedures or with the apparatus was found during the pilot test.

6. Results

The results of the experiment are presented in this chapter. The first section presents the text entry rates with all the techniques used in the experiment. Section 6.2 covers accuracy in terms of error rates with the different techniques. The third section presents findings related to the tilting part of the experiment with the Tilt keyboard method. The final section and its two subsections present the results from the questionnaire as well as the answers gathered from the interview.

6.1. Text entry speeds

The participants of the tilting keyboard experiment wrote a total of 1,140 phrases. 10 phrases were written with each technique per session. For sessions 1-5 this resulted in 30 phrases. Ten more phrases were entered in session 6 because of the additional testing of the QWERTY UK layout.

The means of the results for every text entry method used are shown in Figure 19. For the QWERTY method, the first and last session means were 11.8 WPM and 16.4 WPM, respectively. Overall, the mean improved with QWERTY, with the last session having 39 percent higher WPM than the first.

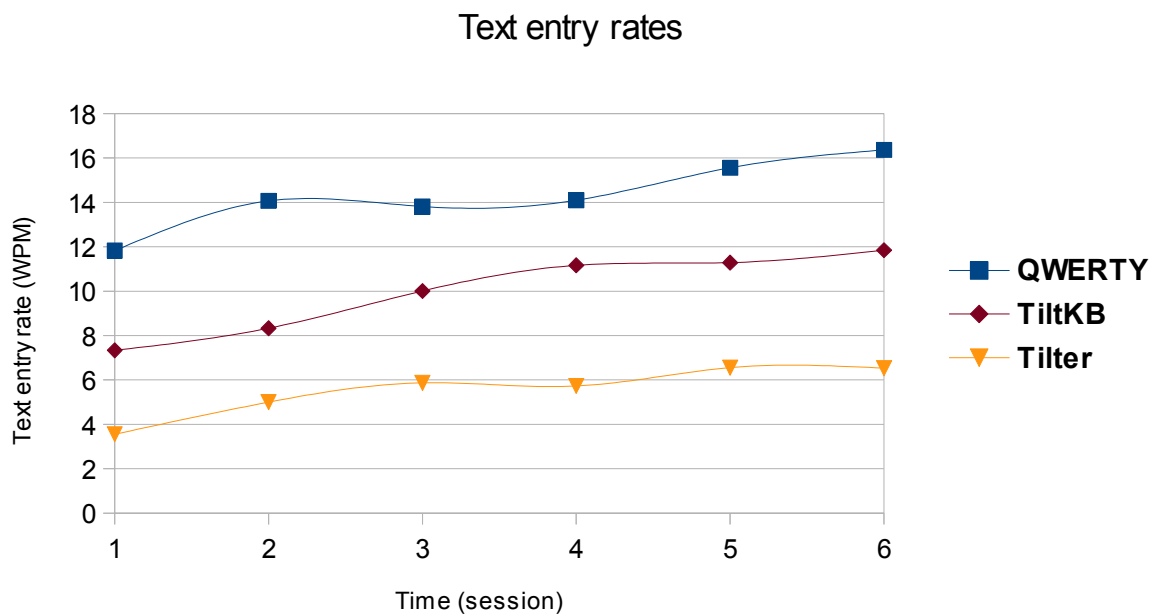


Figure 19: Mean figures showing participants results in words per minute from the first to the sixth session for the QWERTY, Tilt keyboard and Tilter methods.

For the Tilt keyboard method, the first and last session means were 7.3 WPM and 11.9 WPM, respectively. Overall, the mean improved with the Tilt keyboard, with the last session having 63 percent higher WPM than the first.

For the Tilter method, the first and last session means were 3.6 WPM and 6.5 WPM, respectively. Overall, the mean improved with Tilter, with the last session having with 80 percent higher WPM than the first session.

QWERTY was the fastest of the three methods tested, having 38 percent higher text entry rate than the Tilt keyboard and 152 percent higher than Tilter in the last session. The Tilt keyboard had 83 percent higher text entry rate than Tilter in the last session. The QWERTY UK method that was tested in the last session had a mean text entry rate of 17.8 WPM. Therefore, participants had 9 percent faster text entry rate with the QWERTY UK method than with the QWERTY Nordic method in the last session.

The method had a statistically significant ($F(2,10)=54$, $p<0.001$) effect on text entry rate. The effect of the session (i.e. training) was also statistically significant ($F(5,25)=42$, $p<0.001$) as well as the interaction of method and session ($F(10,50)=3.4$, $p=0.002$). All the pairwise comparisons between the methods were also statistically significant (QWERTY and Tilt keyboard $p=0.015$, QWERTY and Tilter $p=0.001$ and Tilt keyboard and Tilter $p=0.005$).

6.2. Accuracy

The means of the results for errors rates are shown in Figure 20. The method had a statistically significant ($F(2,10)=29$, $p<0.001$) effect on error rate. The Tilt keyboard had the lowest error rate, ranging from 7.2 percent (session 1) to 3.3 percent (session 5). QWERTY was next with error rates of 5.7-13 percent and Tilter had clearly most errors, between 15 and 32 percent. The effects of the session (i.e. training) was also statistically significant ($F(5,25)=23$, $p<0.001$) and also the interaction of method and session ($F(10,50)=5.4$, $p<0.001$).

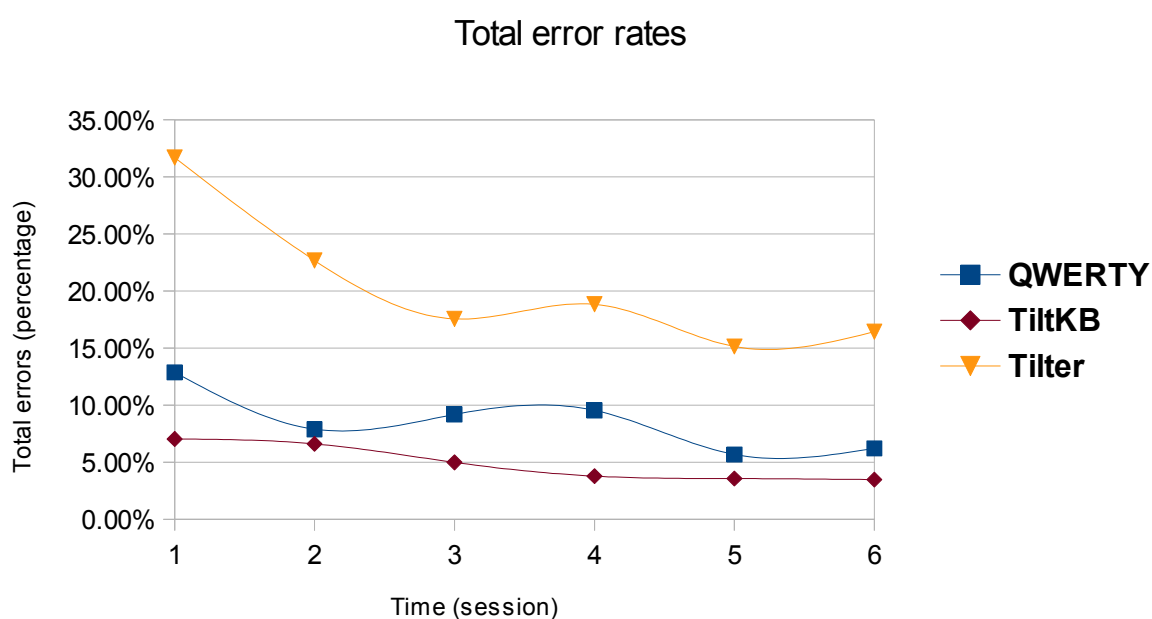


Figure 20: Figures showing individual participants' results in total error rates from the first to the sixth session for all the methods used.

Pairwise comparisons between the methods in error rates gave following results. The difference in error rates between QWERTY and Tilt keyboard was not statistically significant ($p=0.148$), but was statistically significant between QWERTY and Tilter ($p=0.025$). Also, the difference between Tilt keyboard and Tilter was statistically significant ($p=0.019$).

6.3. Tilt efficiency rates with the Tilt keyboard

The minimum tilt count to enter the 500 phrases in the TEMA software is 6,307 tilts. This number includes the need to input enter at the end of each phrase. The total number of characters in these phrases is 14,310 and the average minimum tilts needed per phrase is 12.6. The mean amount of characters per tilt for the test corpus is 2.5, without counting the enter at the end of each phrase.

Figure 21 shows the mean number of characters per tilt efficiency rates for all the participants and sessions. This figure does not take into account the need to input enter at the end of each phrase. In the first session the mean number of characters per tilt efficiency rate was 50 percent and in the last session 66 percent. The curve starts rising in session 3 and is at 63 percent efficiency. The curve peaks at session 4, having 72 percent efficiency.

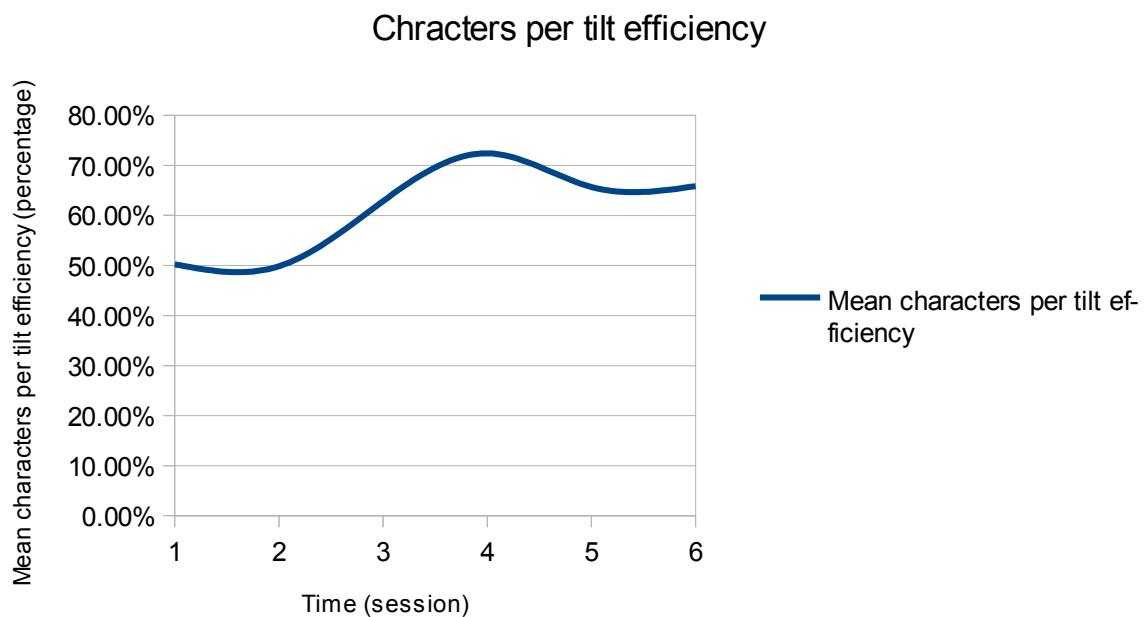


Figure 21: Mean characters per tilt efficiency for all the participants and sessions.

Figure 22 shows characters per tilt efficiency for individual participants. The data shows that characters per tilt efficiency ranged from 22 percent to 66 percent in the first session. In the last session, characters per tilt efficiency was between 38 and 93 percent.

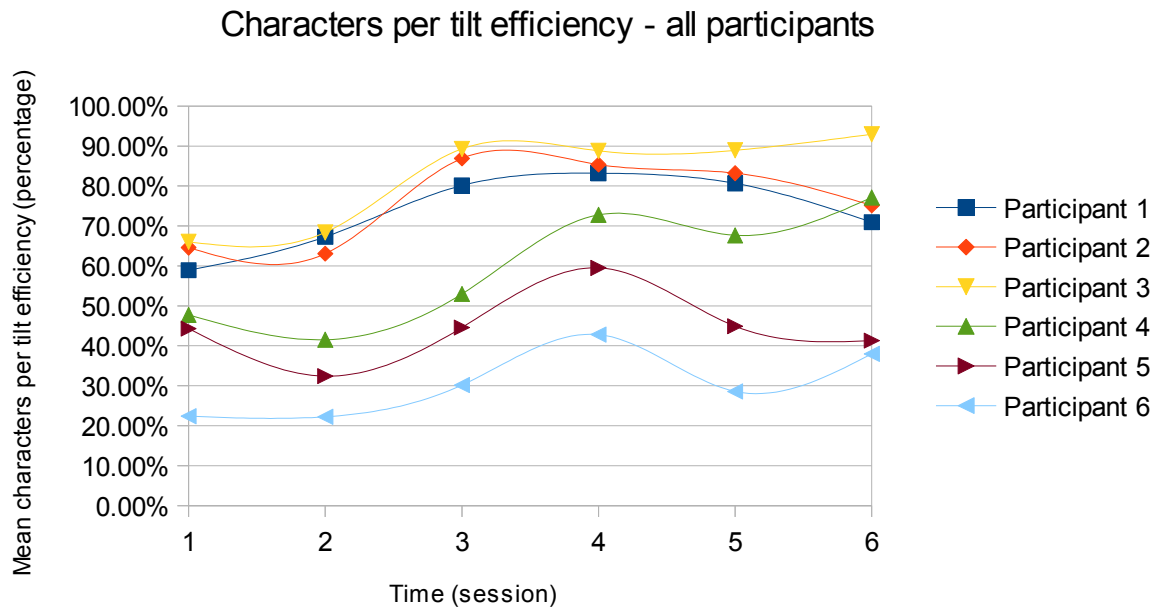


Figure 22: Mean characters per tilt efficiency (see 5.4.) per session for all the participants and sessions.

Participants 5 and 6 had significantly lower characters per tilt efficiency rates than participants 1, 2, 3 and 4, with the minimum difference being 30 percent at the end of the experiment.

6.4. Subjective satisfaction of participants

This section contains two different methods that were used to evaluate the usability and user satisfaction of the tested methods. The first part contains the results of the questionnaire and the second part contains the results of interviews conducted in the first and last sessions.

6.4.1. Questionnaire results

In the first and last session, the participants filled the NASA-TLX based questionnaire (Appendix 3) after using each method. As mentioned earlier in section 5.4.1., the questionnaire used a 21-point scale. A lower number indicated a "better" review for each factor, e.g. ranging from very low to very high for the mental demand and very fast to very low on speed of use.

The mean figures from the questionnaire for the QWERTY method can be seen in Figure 23. In the first session mean Mental demand in the 21-point scale was rated 8.2, Frustration 9.8, Task success 5.7, Speed of use 10.7, Ease of use 9.5 and Learnability as 4.8. The mean rating for all the factors was 8.1.

In the last session, each factors mean were 5.8 for the Mental demand, 8 for Frustration, 6.2 for Task success, 7.5 for Speed of use, 7 for Ease of use and finally 2.5 for Learnability. Mean across the ratings was 6.2.

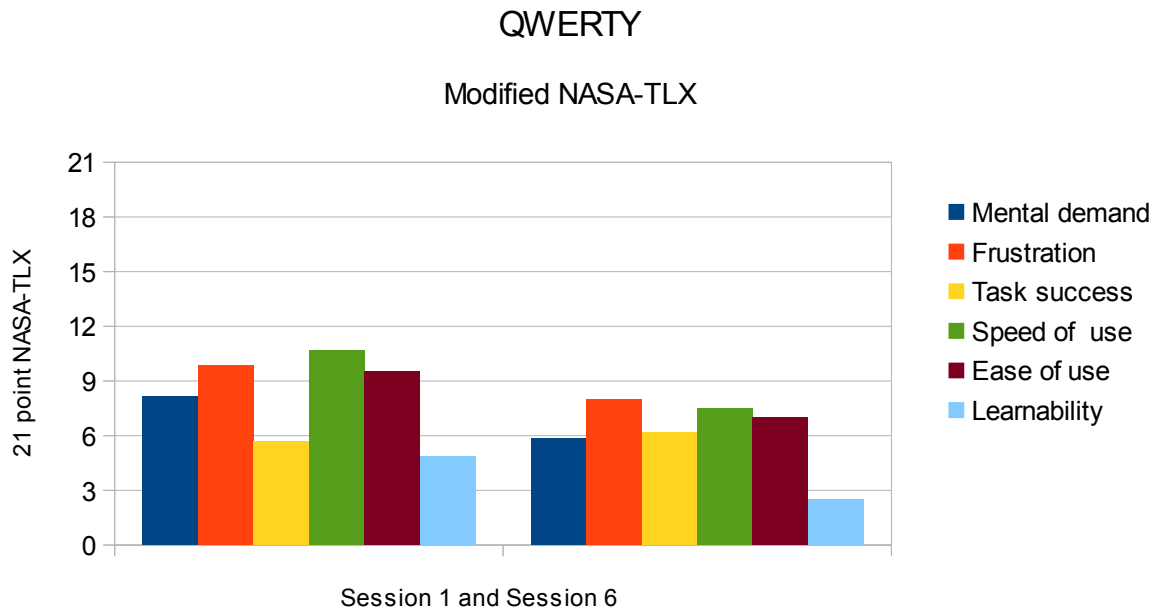


Figure 23: Mean ratings for the NASA-TLX based questionnaire used to evaluate the QWERTY method in the first and last sessions. A lower number is better.

Figure 24 shows the mean figures of the questionnaire for the Tilt keyboard method between the first and last session. In the first session the mean Mental demand for the Tilt keyboard method was rated as 13.7, Frustration as 8.5, Task success as 6, Speed of use as 11, Ease of use as 10 and Learnability as 8.2. The average rating for the factors in the first session was 9.6.

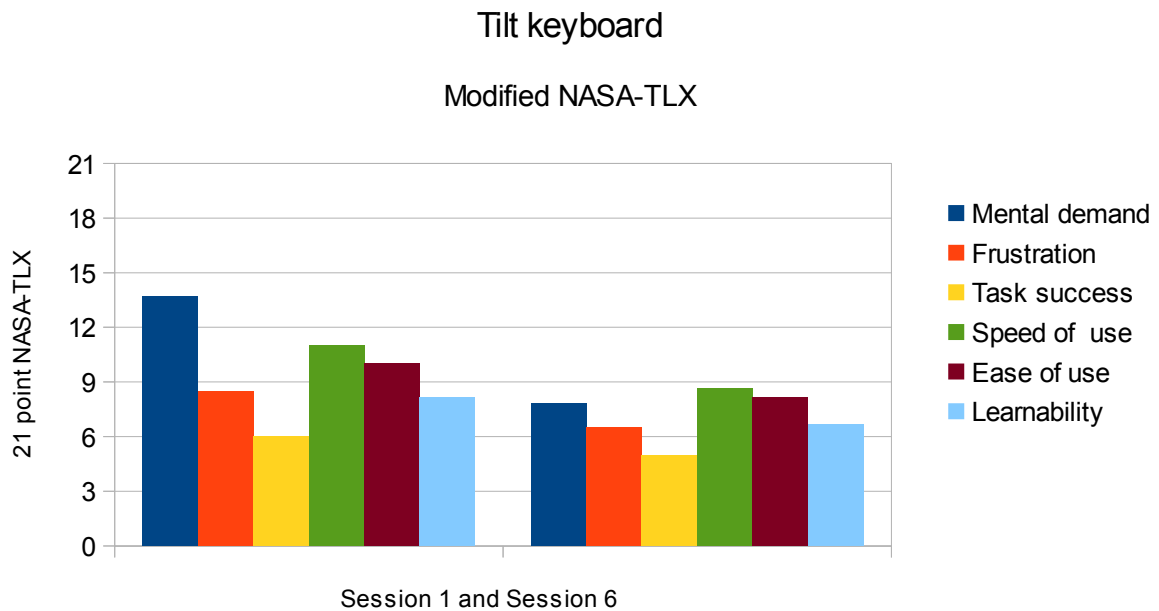


Figure 24: Mean ratings for the NASA-TLX based questionnaire used to evaluate the Tilt keyboard method in the first and last sessions. Lower number is better.

In the last session, the given ratings for the factors were 7.8 for Mental demand, 6.8 for Frustration, 5 for Task success, 8.7 for Speed of use, 8.2 for Ease of use and 6.7 for Learnability. The mean Figure for factors in the last session was 7.1.

Figure 25 shows mean ratings for the factors for the Tilter method. Mental demand was rated as being 16.7, Frustration as 15.3, Task success as 11, Speed of use as 16, Ease of use as 15 and Learnability as being 11.2 in the first session. The mean for all the factors in the first session was 14.2.

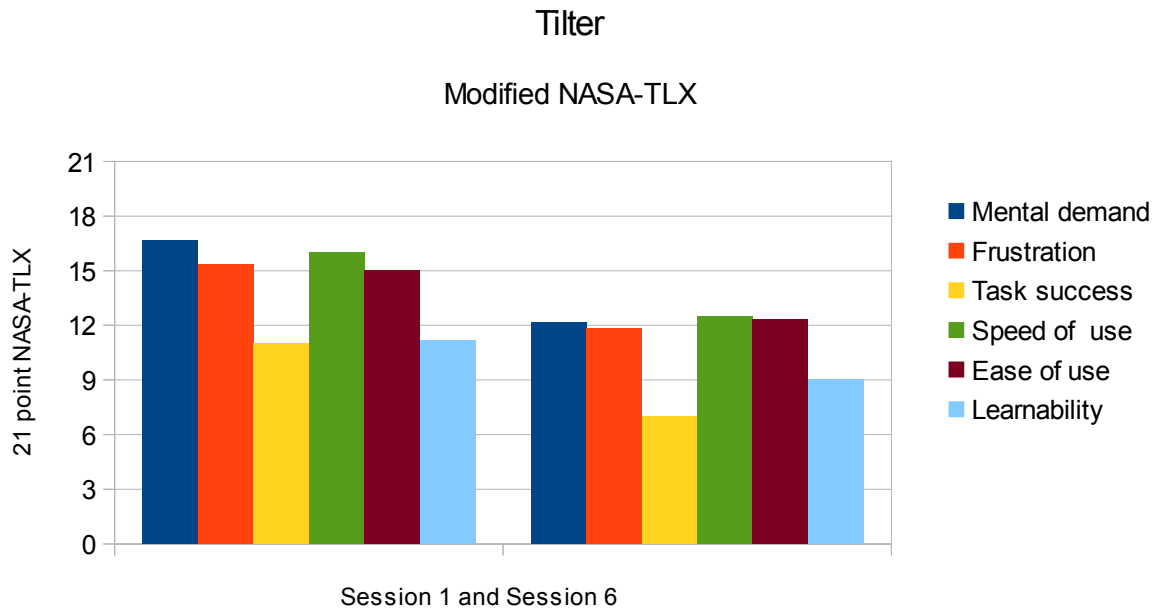


Figure 25: Mean ratings for the NASA-TLX based questionnaire used to evaluate the Tilter method in the first and last sessions. Lower number is better.

In the last session, the mean ratings were 12.2 for the Mental demand, 11.8 for Frustration, 7 for Task success, 12.5 for Speed of use, 12.3 for Ease of use and 9 for Learnability. The mean for the factors in the last session was 10.8.

Overall the QWERTY method had the lowest mean figures across all the factors both in the first and last session. However, the difference to the Tilt keyboard was only 1.5 points in the first and 1.1 points in the last session. Tilter had clearly higher score in the first and in the last session. All the individual factors with all the methods had lower mean score in the last session than in the first session, except the task success for the QWERTY method. Most of the time the improvement was between 1-3 points in the used scale. The biggest improvement for a single factor was for Mental demand of the Tilt keyboard, which improved total of 5.9 points between the first and the last session.

6.4.2. Interview

At the end of the first and last session a short interview was conducted. In this interview, participants were asked to answer three questions The questions were 1) which method did you like best 2) which method felt fastest and 3) what would you use. All the questions were asked in perspective of small device, like the one used in the experiment, and with one hand only use.

The participants' answers can be seen in Table 5. If a participant was not sure of their answer and gave two answers, both are shown in the table in the order they were given. Answers to the main questions were divided quite evenly between the QWERTY and the Tilt keyboard methods. The absence of Tilter can be clearly seen as none of the participants answered it to interview questions.

	Which method did you like best		Which method felt fastest		What would you use		Most common answer
Session	1	6	1	6	1	6	
Participant 1	QWERTY	Tiltkb	Tiltkb	Tiltkb	Tiltkb	Tiltkb	Tiltkb
Participant 2	Tiltkb	QWERTY	Tiltkb QWERTY	Tiltkb QWERTY	QWERTY	QWERTY	QWERTY
Participant 3	QWERTY Tiltkb	Tiltkb	QWERTY	QWERTY	QWERTY	Tiltkb	QWERTY
Participant 4	Tiltkb	Tiltkb	QWERTY Tiltkb	QWERTY	Tiltkb	Tiltkb	Tiltkb
Participant 5	QWERTY	QWERTY	QWERTY	QWERTY	QWERTY	QWERTY	QWERTY
Participant 6	Tiltkb QWERTY	QWERTY	QWERTY	QWERTY	Tiltkb	QWERTY	QWERTY
Most common answer	QWERTY Tiltkb	QWERTY Tiltkb	QWERTY	QWERTY	QWERTY Tiltkb	QWERTY Tiltkb	

Table 5: Answers from all the participants to the three main questions asked in the interview at the end of the first and last sessions

In the first session for the "Which method did you like best" question, QWERTY and the Tilt keyboard were most common answers. In the sixth session, answers to the same questions were divided three for the Tilt keyboard and three for QWERTY. Some of the participants changed their answers between sessions.

The second question was "Which method felt fastest". QWERTY was the most common answer in the first and last session, although the Tilt keyboard was mentioned by few participants.

The third question was "What would you use". The answers were divided evenly between the QWERTY and Tilt keyboard, each getting three answers both in the first and last session. The answers did, however, change between the sessions for individual participants as participant 3 answered QWERTY in the first and Tilt keyboard in the last session and participant 6 answered Tilt keyboard in the first and QWERTY in the last session.

7. Discussion and conclusions

In this chapter the results of the experiment are discussed and briefly compared against some of the other text entry techniques. First five sections focus on the findings regarding the results of the experiment and also on the comments gathered from the participants during the experiment. The last section contains conclusions and future work ideas for tilting-based interaction.

7.1. Text entry speed development

The mean text entry rate for all the tested methods did not offer any surprises at the start of the experiment. The reference QWERTY method was clearly the fastest and Tilter the slowest of the three, while the Tilt keyboard was between the two. All the tested methods improved as sessions progressed, as was expected. However, the difference between the methods was bigger at the end of the testing than was expected. The reference QWERTY method improved clearly between the first and last session, having a total of 39 percent higher text entry speed, which was more than was expected. The likely reason for this is the fact that the testing was done with one hand only and with a small screen device and not the way the participants were accustomed to write on their own devices. Also, the reference method was not the primary method for text entry for two out of six participants, and therefore their improvements are easy to understand.

Although the Tilt keyboard improved more, having 63 percent higher text entry speed in the last than in the first session, it was clearly slower than the reference QWERTY method. The Tilt keyboard's text entry rate in the last session is almost identical to that of the QWERTY method in the first session. The Tilter keyboard was clearly slowest of the three. Although the technique had the biggest improvement between the first and the last session, a total increase of 80 percent, the text entry rate was still slower in the last session than that of the Tilt keyboard in the first session.

The highest recorded text entry rate for any session with QWERTY was achieved by participant 4 in the fifth session, 20.6 WPM. This was also the highest recorded text entry rate for any technique in the experiment with the device used in the final experiment. Also, participant 2 and participant 6 had high text entry rates with the QWERTY method. Participants 2 and 4, wrote a lot with their own mobile devices, "Hundreds of words per day" and "Clearly more than ten, but less than a hundred", respectively. Also, their primary choice of text entry technique was virtual QWERTY on their own devices (see Table 2). These markings seems to have clear connection to their high text entry rates with the technique, even though they were not familiar with the small device or the one hand only used in the experiment.

The text entry rate with the QWERTY keyboard was higher than was expected under the circumstances. The use of a small screen size phone was chosen so that the problems of the full

QWERTY on a small screen would be seen. From the results it can be clearly seen that the text entry rate of the QWERTY method was still similar that shown in the earlier experiments with the bigger screen sizes and two hand use (see Table 1), thus the small screen did not have a significant effect on the writing process. Even so, it is clear that at some point the screen can be too small to be used to enter text with the QWERTY layout keyboard. The results of the study indicates that this size is likely to be under three inches of screen size.

There were some noticeable deviations from the group's mean text entry rates by individual participants. Participant 2 was clearly the fastest within the participant group with both of the novel techniques. Participant 2 was familiar with almost all the techniques that were asked about in the background questionnaire, except the QWERTY physical keyboard, which he had not used (see Table 2). Any clear reason for the good results of participant 2 is unknown. However, he seemed to master the tilting part of the technique quite well and found good neutral position easily. This might partly explain his good results. Participant 2 was also the only left-handed participant in the experiment. However, he used the technique with his right hand, as mentioned earlier.

Participant 5 scored the lowest rates within all the tested methods. His text entry speed with the methods were the slowest in both the first and the last session. In the background information he marked the "few words per day" for the average use and on several occasion mentioned that he did not like to write with his mobile phone. The method he used on his own phone was virtual QWERTY (see Table 2). Also, the problems he had with the tilting-part of the Tilt keyboard might affect his results, which we will discuss in detail in section 7.3.

As the test sentences were in English, this might affect the scores because most of the participants were native Finns. Participant 1 was not a native Finnish speaker. However, only participant 3 mentioned the need to recheck the spelling of a particular word because of the English language and said that it probably affected her writing performance. She was also the only one to mark her English skill to be "good". Others, not including participant 1, marked "excellent" (see Table 2).

Isokoski and Linden [2004] have studied how the results of typing foreign and native language differs even when the English skill is classified as being "good". Their group had 16 native Finnish participants with good English skills and they wrote the test sentences in both languages. According to their results, Finns writing English were 16 percent slower than when writing Finnish. Even though this 16 percent difference can not be said to be a "universal conversion factor" because of several issues which can affect the scores (i.e. the overall speed of text entry, cognitive load of text entry method) it can be said that it is likely that Finnish participants would have written somewhat faster if the sentence provided had been in Finnish.

7.2. Error rate variance

It was expected that there would be some differences in the error rates between the methods used. It was also expected that error rates would decrease as participants would become familiar with the techniques. In the case of Tilter, it was expected that the error rate would be significantly higher in the first sessions than the other two methods, mainly because of the error proneness of the technique. This, in fact, was the case. Also, even though 4 out of 6 participants marked QWERTY as their primary text entry method for their own phone (see Table 2), it was expected that there would be a relatively high error rate at least in the first session. The reason for this was the one hand only use in the experiment, and also the small display size of the used phone.

Data shows that the error rate with the QWERTY method decreases for almost all the participants as the sessions went on. Participant 6 was the only one who did not improve clearly from the first to last session as the error rate with the QWERTY method was 5 percent in the first and 4 percent in the last session.

As for Tilt keyboard, it was expected that the error rates would be lower than with other tested methods because the bigger keys would be easier to hit. This was the case with most of the participants, but not all. When comparing progress between the Tilt keyboard and QWERTY methods, we can see that participants 5 and 6 had similar error rates with the Tilt keyboard and the QWERTY method throughout the six sessions. Participants 1, 2 and 3 always had a lower total error rate with the Tilt keyboard than the QWERTY method. It is hard to find any clear reason why some of the participants were unable to utilize the bigger keys of the Tilt keyboard to their advantage in terms of the total error rate. One reason for this might be the difficulties they had with the tilting part of the technique, which is discussed in the next section.

Participants 2, 4 and 5 used Tilter in the experiment before the Tilt keyboard method, which has the same keyboard layout. It was expected that those participants who would first use Tilter from the two novel techniques, would clearly have higher error rates. The expected reason for this was that the keyboard layouts were not familiar and erroneous page selection automatically means incorrect character. This was not the case as the participants 2 and 4 had the two lowest total error rates with the Tilter method in the first session. The data shows that participants 1, 5 and 6 had clearly higher error rates (between 37 and 45 percent in the first session) than participants 2, 3 and 4 (between 20 and 24 percent). The explanation for this is unclear, but it might be that they did not remember the instructions correctly and thus did not utilize the fact that the most common characters can be found from the right side and instead tried to find them more randomly. However, the relatively small test group makes it difficult to draw any firm conclusions on the matter. In general, the total error rates for the Tilter method were significantly higher than those of the other two methods.

Overall, the total error rates for all the techniques used in the experiment showed an improvement between the sessions. With all three techniques, the total error rate showed around

50% improvement between the first and last session. Tilter had significantly more errors than the two other methods, as was expected. The Tilt keyboard had the lowest total error rate, 3% in the last session, while the QWERTY keyboard had 6%. There is not much to be concluded about the error rates between the techniques. The Tilter's high error rate was to be expected, because the wrong page selection automatically means an erroneous character. The error difference between the Tilt keyboard and the QWERTY keyboard is likely to be explained by the size of the keys

7.3. Problems regarding tilting

As mentioned in the section 6.3., participants 5 and 6 had clearly lower characters per tilt efficiency rates than other participants. One reason that partly explains the poor scores of participant 5 was the fact that he forgot to reset the neutral position on many occasions at the beginning of each test phrase. Participant 6 also forgot to reset the neutral position, and even when it was reset, the phone was used in a position that did not reflect the actual neutral position of the user. This clearly resulted in unintentional keyboard page changes. Other participants did not seem to have this problem. Participant 6 was the only one in the experiment who, during the experiment, stopped the actual writing in the middle of a test phrase to test the neutral position and the page changes, with which he was not satisfied. The problems they had might also partly explain the higher error rate and lower text entry rates of participants 5 and 6, which were mentioned in the earlier sections.

While observing the experiment, these participants had clearly more trouble using the tilting aspect of the text entry technique in general. However, this did not seem to have an effect when using the Tilter method, as the neutral position is not set the same way it is for the Tilt keyboard. Any clear reasons for this are unknown, but one possible explanation might be the natural preferred position of the phone for individual participants. It was noticed that the participants who preferred to use the mobile device at an angle of around 45 degrees towards themselves while writing had more trouble finding a correct or suitable neutral position in the Tilt keyboard technique. Users who preferred to keep their phone more on a horizontal plane did not seem to have as many problems with this aspect of the technique.

7.4. Questionnaire behaviour

All the factors measured using the NASA-TLX based questionnaire improved or remained the same on the 21-point scale used in the questionnaire, with all three techniques between the first and last sessions. The only exception was Task success with the QWERTY method, which was rated 5.67 in the first session and 6.17 in the last session on the 21-point scale.

The biggest improvement when taking the mean of all the factors measured happened with Tilter, total of 3.4 points in the 21-point scale. The next was the Tilt keyboard: from 9.6 to 7.1, an

improvement of 2.4 points. The last was QWERTY: from 8.1 to 6.2, an improvement of 1.9 points in the used scale from the first to last session. It is hard to assess the usability of methods on the basis of these numbers, especially because the pairwise comparison of each contribution factor was not conducted in this experiment like mentioned earlier in section 5.4.1. However, it is fairly safe to say that all the methods were easier to use in the last session than in the first session.

The individual participants' ratings followed mean ratings closely and most of the time participants gave 1-3 points better or same score for each factor on every method, between the first and the last sessions. The most noticeable improvements were given by participant 3, who gave clearly better scores for all the factors for the Tilt keyboard in the last session. The biggest improvements of these were given to Mental demand, from 16 to 2, and to Speed of use, from 16 to 6, in the 21-point scale.

On many occasions the participants also gave a poorer score to some factor in the last session. Most of the time the given numbers were not significant and the difference was 1-3 points on the used scale. Most noticeable of these were given by participant 6. He gave QWERTY a worse rating in the last session in Mental demand, Frustration and Task success than in the first session. Task success increased from 6 to 12, even though the method was not used by the participant on his own mobile phone and he improved significantly in terms of text entry speed. He also gave the Tilt keyboard a worse score on almost every factor in the questionnaire. This was the case even though he improved with the technique between the first and the last session in terms of text entry rate. It is possible that the problems he had with the tilting part of the technique affected the judgement of the method in the questionnaire.

7.5. Comments by the participants

The participants also had the opportunity to give comments on the used methods at the end of each session. The most common comments are discussed here.

The QWERTY method was mostly criticized for the small size of the keys and the difficulty of use with one hand only, which often resulted in wrong characters being entered. However, the participants also commented that the keys were relatively easy to hit and that they had expected to make more errors. It was also commented that even though the participants experienced a lot of problems and frustration in the early sessions, writing with the technique became easier as the sessions progressed. Still, the amount of errors and the "need to aim and concentrate" were frustrating according to participants. Some of the participants also commented that it was quite hard to hit certain keys with the thumb. The positive comments for the QWERTY method were mainly about the familiarity of the technique and that even when there was an error, correcting it happened almost automatically.

In general, the participants gave positive feedback about the Tilt keyboard method. Most of the participants commented that writing with it was easy when the layouts were familiar and the neutral

position in the correct place. At the end of the experiment, all the participants said that the layouts were getting familiar, but that more sessions would be needed to completely familiarize themselves with the layouts.

Two participants commented that some characters, especially ones from the upper page, were hard to remember as they were used so rarely. Some participants commented that the technique would be good in situations where the surrounding is moving, e.g. in a bus or when writing with one hand only like in the experiment. Most of the participants said that writing with the technique would likely be faster with two hands, although one participant did not like the idea of using the technique with two hands at all.

Most participants commented that the keys were big enough, but participant 5 said that they are still too small to avoid errors. Most participants had trouble using the tilting part of the technique in the first two sessions. Participant 2 was the only one who seemed to master it from the start. Others tilted more than was needed or had trouble setting a good neutral position at first. Participants 4 and 6 also commented that the tilting is too sensitive for them while others did not have problems with the sensitivity. Participant 4 also commented that when reaching for a key, the page sometimes changed unintentionally. None of the participants had problems with their wrist when writing with the Tilt keyboard, and a possible sore wrist in long writing session was mentioned only once by participant 2.

As for the Tilter method, most users were quite frustrated with the use of the technique in the first sessions, which can be seen from the high frustration ratings given in the questionnaire. The participants did not like the fact that the layout page could not be changed after the page had been selected. This was felt as frustrating at least in the first sessions, when the layouts were not familiar for the participants. It was also commented that the technique is somewhat mentally demanding. The reason mentioned for this was the multiple phases needed to enter a character while trying to keep in mind what key is needed next and on what page it is located. All the participants complained that the technique is straining for the wrist. The most difficult thing from the wrist straining perspective was when there was a need to enter a character that was located on the same side of an individual page than the page itself. For example, when entering the character 'd', the user first tilts to the right to select page and then again to the right to select the character. This resulted an extreme position for the wrist. The effect was even more troublesome when there were two d's in a row and the user often did not realize to move phone back to a horizontal plane between the characters. The position of the volume up button, which worked as an activation key for the tilting mechanism (see 5.3.), also gathered comments, and most of the participants found it to be bad. The screen lock button was located on the other side of the phone and this caused problems for some of the participants.

Some of the participants said that the method was cognitively tiring. For example, three participants had a similar problem when predicting how to enter the next needed character. For example when they were writing word "cat" and were entering the first character 'c', which is located in the right page, they saw the next character 'a' on the middle left of that same page. Next

when they wanted to enter the character 'a' they tilted directly to the left and therefore entered to the wrong page. These kinds of errors decreased as the sessions progressed and some participants commented that instead of selecting a page and then a character, the process became more automatic: selecting the character 'a', for example, became a fast tilt movement from first to the right and then to the left.

7.6. Additional remarks about the results

The comparison of the text entry rate of the tilt keyboard method to the other methods in the previous studies shows that the Multi-tap method has around the same recorded text entry rates, 7.2 – 15.5 WPM (see Table 1), but the T9 method is faster (9.1 – 20.4 WPM). The Tilter method has around the same text entry rate as EdgeWrite in the study conducted by Wobbrock et al [2003]. The text entry rate of the virtual QWERTY keyboard was about the same as in the previous studies by Költringer and Grechenig [2004] and MacKenzie *et al.* [2009], between 13.6 WPM and 18.5 WPM in the last session, even though the phone used in this experiment was relatively small and was written using one hand only.

When comparing the text entry rate of other techniques that utilize tilting as part of the input of text, the Tilt keyboard shows to have same similar text entry rates. The experiment by Wigdor and Balakrishnan [2003] with their TiltText method shows to have text entry rates at the end of the experiment 13.6 WPM. SHRIMP by Wang *et al.*, [2010] which combined dictionary-based disambiguation and tilting showed a text entry rate of 12.1 WPM. The Tilter method shows to be clearly slower than either of these methods, which was to be expected.

Some factors should be considered when evaluating the results of the experiment. The use of one hand only has an effect on the results, at least with the Tilt keyboard and the QWERTY methods. None of the participants had previous experience writing with one hand only and most struggled to reach certain keys comfortably. While the keys on the Tilt keyboard are bigger and thus in general easier to reach and press, the need to reach sometimes resulted in unintentional tilting and the wrong layout page. The effect on the results is difficult to determine, but it is likely that the results would have some differences if the techniques would be used with a bigger device and/or with both hands. The Tilter method would probably not benefit from the use of both hands, although one participant mentioned that two hands would result in a steadier tilting experience. Also, the techniques were tested in ideal conditions. For example reflections (e.g. from the sky or lights) are likely to have an affect when using the tilting-based techniques in outdoors.

As mentioned in section 5.5., the participants were instructed to correct errors as they occurred, but not to correct errors they noticed later that were located in the earlier parts of the current phrase. In many cases, however, participants did not follow this instruction, sometimes deleting many words to correct one particular typo so that the entered phrase would be perfect. This happened with

all the techniques and thus can be assumed to have equally affected the text entry rate of all methods.

In the case of Tilter, the participants pressed the wrong button on several occasion, volume down, when their intention was to press volume up and enter a character. Sometimes this resulted in multiple words being deleted by accident before the participant realised that the pressed button was the wrong one. This is clearly a defect of the method design and therefore these results are taken into account. Some of the participants commented that the buttons are not located in optimal places for the technique and that this affected their results somewhat.

The participants gave generally quite positive remarks about the usage of the novel techniques in the interview and when commenting freely on the techniques. For example the Tilt keyboard was mentioned four times in the first interview question, "which method you liked best", even though participants had only written 10 phrases with it at that time and were still novice users with the technique. This was the case even though sometimes the technique was not felt to be the fastest of the three, which was the subject of the second question of the interview. One reason for this might be that the participants did not find the QWERTY method easy to use with the small device and one hand only. As mentioned earlier, the participants answered the questions from the perspective of a small screen size and one hand use in mind, like the device used in the testing. The third question "what would you use" likely favours this setup, and the Tilt keyboard is not likely to get high remarks in terms of normal usage (e.g. with the use of a bigger phone and two hands).

7.7. Conclusions and future work

Tilting as an input method has been studied, as part of text entry techniques as well as for other applications. Even so, practical implementations are few in number. The bandwidth between man and machine can be increased by adding tilting as part of an input modality. However, as the experiment here shows, adding a multimodal approach does not always increase the performance of any particular technique.

The main goal of the thesis was to evaluate the tilting-based text entry techniques from many aspects. Even though the results clearly show that both novel techniques were slower than the chosen reference method in this setup, they also show, that tilting can be used as part of a text entry technique on smartphones, and can even be the preferred input method in some cases. All the participants were able to write using both novel tilting techniques and their writing clearly improved during the experiment.

Many questions regarding the use of tilting as part of text entry remain, however, which could be studied in follow-up experiments, mainly how well the techniques would do when used with both hands or with smaller/bigger devices. These techniques could also be used on devices with screen sizes so small that the use of a normal QWERTY keyboard is not feasible. Also, the Tilter method could be used, in theory, without a screen if the layouts were memorized by the user. Even

though the Tilter and the Tilt keyboards worked quite well during the experiment, they could benefit from some re-programming and designing. Especially the setting of the neutral position, which caused some participants problems in the experiment, could be redesigned. Also, the Tilter method could be redesigned so that it would not return to the start page after selecting a character, which seemed to cause problems for some of the participants, at least at the start of the experiment.

If the use of tilting as part of text entry methods increases, standardized analysis tools are needed, much like the Text Entry Metrics for Android is for general performance measurement. Analysis tools like these could record all the tilting events and calculate the most essential statistics automatically. This would be beneficial for all parties as the new methods do not need to create a way to record their tilt events. Also, comparing the results between methods of different authors would become easier.

As the results show, text entry with tilting is clearly possible and therefore could be used in more suitable situations. These might be, for example, as part of a text entry on devices with extremely small screen sizes, like smartwatches that have entered the market recently. These devices are usually not intended for text entry as the screen is too small for almost any touch based text entry method. Also, tilting has potential in menu browsing, which has been studied earlier by Turunen *et al.*, [2009], and could be utilized more in situations where browsing menus is hard or impossible by traditional means. Another area where tilting could prove useful is when all the information cannot fit to a small screen at once and scrolling or panning is necessary, like when browsing a map. Tilting could be used as a way to scroll the screen in different directions with one hand only. Further research into the feasibility of this is needed.

Overall I feel that tilting as an input method has potential that is yet to be utilized. Therefore, future studies are needed that would eventually lead to a practical implementation and increased bandwidth between man and machine.

References

- [Algoritmy.net, 2013] Algoritmy.net – Letter frequency (English), <http://en.algoritmy.net/article/40379/Letter-frequency-English> (Checked on December 1st, 2013).
- [Apple-History, 1996] Apple-history – Newton Message Pad (OMP), <http://www.apple-history.com/nmp> (Checked on December 1st, 2013).
- [BloombergBusinessweek, 2012] BloombergBusinessweek Technology – Before iPhone and Android Came Simon, the First Smartphone, <http://www.businessweek.com/articles/2012-06-29/before-iphone-and-android-came-simon-the-first-smartphone> (Checked on October 27th 2013).
- [Brewster *et al.*, 2007] Brewster, S., Chohan, F., and Brown, L. Tactile feedback for mobile interactions. In: *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 159-162). ACM.
- [Bustamante and Spain, 2008] Bustamante, E. A., and Spain, R. D. Measurement Invariance of the NASA TLX. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 52, No. 19, pp. 1522-1526). SAGE Publications.
- [Butts and Cockburn, 2002] Butts, L., and Cockburn, A. An evaluation of mobile phone text input methods. In: *Australian Computer Science Communications* (Vol. 24, Issue 4, pp. 55-59). Australian Computer Society, Inc..
- [Card *et al.*, 1986] Card, S. k., Moran, T. P., and Newell, A. *The Psychology of Human-Computer Interaction*. CRC Press.
- [Castellucci and MacKenzie, 2011] Castellucci, S. J., and MacKenzie, I. S. Gathering text entry metrics on android devices. In: *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 1507-1512). ACM.

- [Clarkson *et al.*, 2005] Clarkson, E., Clawson, J., Lyons, K., and Starner, T. An empirical study of typing rates on mini-QWERTY keyboards. In: *Proceedings of the CHI 2005 Conference on Human Factors in Computing Systems*, (pp.1288–1291). ACM.
- [Cockburn and Siresena, 2003] Cockburn, A., and Siresena, A. Evaluating mobile text entry with the Fastap keypad. http://ir.canterbury.ac.nz/bitstream/10092/508/1/12589072_fastap.pdf (Checked on December 1st, 2013).
- [Complete TLX Publication List, 2013] NASA-TLX, Complete TLX Publication List, http://humansystems.arc.nasa.gov/groups/TLX/downloads/TLX7_95.xls (Checked on December 1st, 2013).
- [Data-compression.com, 2013] Data-compression.com – Statistical Distributions of English Text, <http://www.data-compression.com/english.html> (Checked on December 1st, 2013).
- [Dearman *et al.*, 2010] Dearman, D., Karlson, A., Meyers, B., and Bederson, B. Multi-modal text entry and selection on a mobile device. In: *Proceedings of Graphics Interface 2010* (pp. 19-26). Canadian Information Processing Society.
- [Dvorak and Dealey, 1936] Dvorak, A., and Dealey, D. L. Typewriter keyboard. U.S. Patent 2,040,248. U.S. Patent Office.
- [Elwenspoek and Wiegierink, 2001] Elwenspoek, M., and Remco J. W. In: *Mechanical Microsensors* (pp. 5-23 and 132-152). Springer.
- [Evreinova *et al.*, 2004] Evreinova, T., Evreinov, G., and Raisamo, R. Four-key text entry for physically challenged people. In: *Adjunct Proceedings of the 8th ERCIM Workshop on User Interfaces for All (UI4ALL '04)*.
- [Fallman, 2002] Fallman, D., Wear Point and Tilt. In: *Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (pp. 293-302). ACM Press.
- [Fitts, 1954] Fitts, P. M. The information capacity of the human motor system in controlling the amplitude of movement. In: *Journal of Experimental Psychology* (Vol. 47, No. 6, pp. 381–391).

- [Gartner, 2009] Gartner Says Worldwide Smartphone Sales Reached Its Lowest Growth Rate With 3.7 Per Cent Increase in Fourth Quarter of 2008, <http://www.gartner.com/newsroom/id/910112> (Checked on December 1st, 2013).
- [Gartner, 2013] Gartner Says Smartphone Sales Grew 46.5 Percent in Second Quarter of 2013 and Exceeded Feature Phone Sales for First Time, <http://www.gartner.com/newsroom/id/2573415> (Checked on December 1st, 2013).
- [Glover et al., 1995] Glover, D., L., King, T., and Kushler, C., A., Reduced Keyboard Disambiguating Computer, Patent number 5818437
- [Gong et al., 2008] Gong, J., Tarasewich, P., and MacKenzie, I. S. Improved word list ordering for text entry on ambiguous keypads. In: *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges* (pp. 152-161). ACM.
- [Hart, 2006] Hart, S. G., NASA-TASK LOAD INDEX (NASA-TLX; 20 YEARS LATER, http://sunburst.usd.edu/~schieber/psyc792/workload/Hart_HFES_20006.pdf (Checked on December 1st, 2013).
- [Hoggan et al., 2008] Hoggan, E., Brewster, S. A., and Johnston, J. Investigating the effectiveness of tactile feedback for mobile touchscreens. In: *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 1573-1582). ACM.
- [Ingmarsson et al., 2004] Ingmarsson, M., Dinka, D., and Zhai, S. TNT: a numeric keypad based text input method. In: *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 639-646). ACM.
- [Isokoski and Linden, 2004] Isokoski, P., and Linden, T. Effect of foreign language on text transcription performance: Finns writing English. In: *Proceedings of the third Nordic conference on Human-computer interaction* (pp. 109-112). ACM.
- [Isokoski and Käki, 2002] Isokoski, P., and Käki, M. Comparison of two touchpad-based methods for numeric entry. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 25-32). ACM.
- [ITU-T, 2001] International Telecommunication Union. ITU-T Recommendation E.161: Arrangement of digits, letters and symbols on a telephones and other devices that can be used for gaining access to a telephone network, <http://www.itu.int/rec/T-REC-E.161-200102-I/en> (Checked on December 1st, 2013).

- [ITU-T, 2013] International Telecommunication Union, ICT Facts and Figures 2013, <http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013.pdf> (Checked on December 1st, 2013).
- [James and Reischel, 2001] James, C. L., and Reischel, K. M. Text input for mobile devices: comparing model prediction to actual performance. In: *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 365-371). ACM.
- [Költringer and Grechenig, 2004] Költringer, T., and Grechenig, T. Comparing the immediate usability of graffiti 2 and virtual keyboard. In: *CHI'04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1175-1178). ACM.
- [Luo and John, 2005] Luo, L., and John, B. E. Predicting task execution time on handheld devices using the keystroke-level model. In: *CHI'05 extended abstracts on Human factors in computing systems* (pp. 1605-1608). ACM.
- [MacKenzie, 1992] MacKenzie, I. S. Fitts' law as a research and design tool in human-computer interaction. In: *Human-computer interaction* (Vol. 7, Issue 1, pp. 91-139). Lawrence Erlbaum Associates.
- [MacKenzie, 2002] MacKenzie, I. S. KSPC (keystrokes per character) as a characteristic of text entry techniques. In: *Human Computer Interaction with Mobile Devices* (pp. 195-210). Springer Berlin Heidelberg.
- [MacKenzie, 2007] MacKenzie, I. S. Evaluation of Text Entry Techniques. In: I. Scott MacKenzie and Kumiko Tanaka-Ishii (eds.), *Text Entry Systems: Mobility, Accessibility, Universality* (pp. 75-101). Morgan Kaufmann.
- [MacKenzie and Zhang, 1999] MacKenzie, I. S., and Zhang, S. X. The design and evaluation of a high-performance soft keyboard. In: *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 25-31). ACM.
- [MacKenzie *et al.*, 2001] MacKenzie, I. S., Kober, H., Smith, D., Jones, T., and Skepner, E. LetterWise: prefix-based disambiguation for mobile text input. In: *Proceedings of the 14th annual ACM symposium on User interface software and technology* (pp. 111-120). ACM.

- [MacKenzie and Soukoreff, 2002] MacKenzie, I. S., and Soukoreff, R. W. Text entry for mobile computing: Models and methods, theory and practice. In: *Human-Computer Interaction* (Vol. 17, Issue 2-3, pp. 147-198). Taylor & Francis Online.
- [MacKenzie and Soukoreff, 2003] MacKenzie, I. S., and Soukoreff, R. W. Phrase sets for evaluating text entry techniques. In: *CHI'03 extended abstracts on Human factors in computing systems* (pp. 754-755). ACM.
- [MacKenzie and Tanaka-Ishii, 2007] MacKenzie, I. S., and Tanaka-Ishii, K. Text Entry Using a Small Number of Buttons. In: I. Scott MacKenzie and Kumiko Tanaka-Ishii (eds.), *Text Entry Systems: Mobility, Accessibility, Universality* (pp. 105-122). Morgan Kaufmann.
- [MacKenzie *et al.*, 2009] MacKenzie, I. S., Lopez, M. H., and Castelluci, S. Text Entry with the Apple iPhone and the Nintendo Wii. <http://www.malchevic.com/papers/MobileHCI2009.pdf> (Checked on April 21st, 2014). Unpublished manuscript.
- [MacKenzie and Teather, 2012] MacKenzie, I. S., and Teather, R. J. FittsTilt: the application of Fitts' law to tilt-based interaction. In: *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design* (pp. 568-577). ACM.
- [Majaranta and R  ih  , 2007] Majaranta, P. and R  ij   K-J., Text Entry by Gaze: Utilizing Eye Tracking. In: I. Scott MacKenzie and Kumiko Tanaka-Ishuu (eds.), *Text Entry Systems: Mobility, Accessibility, Universality* (pp. 175-187). Morgan Kaufmann.
- [Martin, 2008] Martin, D. W. In: *Doing Psychology Experiments*. Cengage Learning.
- [Martin *et al.*, 2009] Martin, B., Isokoski, P., Jayet, F., and Schang, T. Performance of finger-operated soft keyboard with and without offset zoom on the pressed key. In: *Proceedings of the 6th International Conference on Mobile Technology, Application & Systems* (Article No. 59). ACM.
- [Matias *et al.*, 1996] Matias, E., MacKenzie, I. S., and Buxton, W. One-handed touch typing on a QWERTY keyboard. In: *Human-Computer Interaction* (Vol. 11, Issue 1, pp. 1-27). Lawrence Erlbaum Associates.
- [NASA-TLX manual, 1986] Task load index (NASA-TLX) V 1.0 Nasa, <http://humansystems.arc.nasa.gov/groups/TLX/downloads/TLX.pdf> (Checked on December 1st, 2013).

- [Nokia Press, 1996] Nokia Press: Nokia Unveils Worlds first All-In-One Communicator For The Americas, <http://press.nokia.com/1996/09/19/nokia-unveils-world%C2%92s-first-all-in-one-communicator-for-the-americas/> (Checked on December 1st, 2013).
- [Noyes and Bruneau, 2007] Noyes, J. M., and Bruneau, D. P. A self-analysis of the NASA-TLX workload measure. In: *Ergonomics* (Vol. **50**, Issue 4, pp. 514-519). Taylor & Francis Online.
- [Nuance, 2013] Nuance T9 Text Input, <http://www.nuance.com/t9/textinput> (Checked on December 1st, 2013).
- [Pakkanen *et al.*, 2010] Pakkanen, T., Raisamo, R., Salminen, K., and Surakka, V. Haptic numbers: three haptic representation models for numbers on a touch screen phone. In: *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction* (Article No. 35). ACM.
- [Partridge *et al.*, 2002] Partridge, K., Chatterjee, S., Sazawal, V., Borriello, G., and Want, R. TiltType: accelerometer-supported text entry for very small devices. In: *Proceedings of the 15th annual ACM symposium on User interface software and technology* (pp. 201-204). ACM.
- [PCMag, 2013] PCMag encyclopedia – Smartphone, <http://www.pcmag.com/encyclopedia/term/51537/smartphone> (Checked on October 27th, 2013).
- [Rahman *et al.*, 2009] Rahman, M., Gustafson, S., Irani, P., and Subramanian, S. Tilt techniques: investigating the dexterity of wrist-based input. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1943-1952). ACM.
- [Samsung mobile press, 2013] The Samsung Mobile press – Samsung Introduces the GALAXY Mega, <http://www.samsungmobilepress.com/2013/04/11/Samsung-Introduces-the-GALAXY-Mega-1> (Checked on June 1st, 2013).
- [Silfverberg, 2007] Silfverberg, M. Historical Overview of Consumer Text Entry Technologies. In: I. Scott MacKenzie and Kumiko Tanaka-Ishuu (eds.), *Text Entry Systems: Mobility, Accessibility, Universality* (pp. 3-26). Morgan Kaufmann.
- [Smith and Goodwin, 1971] Smith, S. L., and Goodwin, N. C. Alphabetic data entry via the Touch-Tone pad: A comment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, (Vol. **13**, No. 2, pp. 189-190). SAGE Publications.

- [Soukoreff and MacKenzie, 2001] Soukoreff, R. W., and MacKenzie, I. S. Measuring errors in text entry tasks: an application of the Levenshtein string distance statistic. In: *CHI'01 extended abstracts on Human factors in computing systems* (pp. 319-320). ACM.
- [Soukoreff and MacKenzie, 2003] Soukoreff, R. W., and MacKenzie, I. S. Metrics for text entry research: an evaluation of MSD and KSPC, and a new unified error metric. In: *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 113-120). ACM.
- [Soukoreff and MacKenzie, 2004] Soukoreff, R. W., and MacKenzie, I. S. Recent developments in text-entry error rate measurement. In: *CHI'04 extended abstracts on Human factors in computing systems* (pp. 1425-1428). ACM.
- [Stefan Trost Media, 2013] Stefan Trost Media: Syllable Frequency: English, <http://www.sttmedia.com/syllablefrequency-english> (Checked on December 1st, 2013).
- [Stockholmsmartphone.org, 2010] Stockholm smartphone – History, <http://www.stockholmsmartphone.org/history/> (Checked on October 27th, 2013).
- [SwypeInc, 2013] Swype, Nuance, <http://www.swype.com/> (Checked on December 1st, 2013).
- [Tanaka-Ishii, 2007] Tanaka-Ishii, K. Language Models for Text Entry. In: I. Scott MacKenzie and Kumiko Tanaka-Ishuu (eds.), *Text Entry Systems: Mobility, Accessibility, Universality* (pp. 27-45). Morgan Kaufmann.
- [Tappert and Cha, 2007] Tappert, C. C., and Cha, S. English Language Handwriting Recognition Interfaces. In: I. Scott MacKenzie and Kumiko Tanaka-Ishuu (eds.), *Text Entry Systems: Mobility, Accessibility, Universality* (pp. 123-138). Morgan Kaufmann.
- [Textware Solutions, 1999] Fitaly One-Finger Keyboard. <http://fitaly.com/fitaly/fitaly.htm> (Checked on December 1st, 2013).
- [Trewin and Arnott, 2007] Trewin, S. and Arnott, J., Text Entry When Movement is Impaired. In: I. Scott MacKenzie and Kumiko Tanaka-Ishuu (eds.), *Text Entry Systems: Mobility, Accessibility, Universality* (pp. 289-304). Morgan Kaufmann.

- [Turunen *et al.*, 2009] Turunen, M., Melto, A., Hella, J., Heimonen, T., Hakulinen, J., Mäkinen, E., and Soronen, H. User expectations and user experience with different modalities in a mobile phone controlled home entertainment system. In: *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Article No. 31). ACM.
- [Venolia and Neiber, 1994] Venolia, D., and Neiber, F. T-Cube: a fast, self-disclosing pen-based alphabet. In: *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 265-270). ACM.
- [Wang *et al.*, 2006] Wang, J., Zhai, S., and Canny, J. Camera phone based motion sensing: interaction techniques, applications and performance study. In: *Proceedings of the 19th annual ACM symposium on User interface software and technology* (pp. 101-110). ACM.
- [Wang *et al.*, 2010] Wang, J., Zhai, S., and Canny, J. SHRIMP: solving collision and out of vocabulary problems in mobile predictive input with motion gesture. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 15-24). ACM.
- [Ward *et al.*, 2000] Ward, D. J., Blackwell, A. F., and MacKay, D. J. Dasher—a data entry interface using continuous gestures and language models. In: *Proceedings of the 13th annual ACM symposium on User interface software and technology* (pp. 129-137). ACM.
- [Ward and MacKay, 2002] Ward, D. J., and MacKay, D. J. Artificial intelligence: Fast hands-free writing by gaze direction. In: *Nature* (Vol. **418**, p. 838).
- [Wigdor and Balakrishnan, 2003] Wigdor, D., and Balakrishnan, R. TiltText: using tilt for text input to mobile phones. In: *Proceedings of the 16th annual ACM symposium on User interface software and technology* (pp. 81-90). ACM.
- [Wobbrock, 2005] Wobbrock, J. O. A robust design for accessible text entry. In: *ACM SIGACCESS Accessibility and Computing* (Issue 84, pp. 48-51). ACM.
- [Wobbrock, 2007] Wobbrock, J. O. Measures of text entry performance. In: I. Scott MacKenzie and Kumiko Tanaka-Ishii (eds.), *Text Entry Systems: Mobility, Accessibility, Universality* (pp. 47-74). Morgan Kaufmann.
- [Wobbrock *et al.*, 2003] Wobbrock, J. O., Myers, B. A., and Kembel, J. A. EdgeWrite: a stylus-based text entry method designed for high accuracy and stability of motion. In: *Proceedings*

of the 16th annual ACM symposium on User interface software and technology, (pp. 61-70). ACM.

- [Wobbrock et. al, 2006] Wobbrock, J., Myers, B., and Rothrock, B. Few-key text entry revisited: mnemonic gestures on four keys. In: *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 489-492). ACM.
- [Wobbrock and Myers, 2006] Wobbrock, J. O., and Myers, B. A. Analyzing the input stream for character-level errors in unconstrained text entry evaluations. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* (Vol. **13**, Issue 4, pp. 458-489). ACM.
- [Yamada, 1980] Yamada, H. A historical study of typewriters and typing methods: From the position of planning Japanese parallels. In: *Journal of Information Processing* (Vol. **2**, Issue 4, pp. 175-202). CiNii.
- [Zhai et al., 2002] Zhai, S., Hunter, M., and Smith, B. A. Performance optimization of virtual keyboards. In: *Human-Computer Interaction* (Vol. **17**, Issue 2-3, pp. 229-269). Taylor & Francis Online.
- [Zhai and Kristensson, 2003] Zhai, S., and Kristensson, P. O. Shorthand writing on stylus keyboard. In: *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 97-104). ACM.
- [Zhai and Kristensson, 2006] Zhai, S., and Kristensson, P. O. Introduction to ShapeWriting, *IBM Research Report RJ10393 (A0611-006)*

tilting-based text entry, questionnaire

Background information

Age: _____ Gender: _____ Handedness: _____

Education:

- ☐ Primary school
- ☐ Vocational school
- ☐ Gymnasium / Grammar school / High school
- ☐ Bachelor's degree
- ☐ Master's degree / Ph. D.

English language skills:

- ☐ Weak
- ☐ Satisfactory
- ☐ Good
- ☐ Excellent
- ☐ Native speaker

User experience

1. How would you describe yourself as an information technology user in general?

- ☐ Experienced user, I have used many different devices and I am comfortable using them.
- ☐ I manage, but I don't have that much experience with different devices.
- ☐ I have often need help for the problems different devices and applications cause for me.

2. What is your primary mobile device (phone/tablet), give manufacturer, name, model and screen size:

3. Give an estimate how many words you typically write with your mobile device on a average day?

- ☐ Over a thousand words per day
- ☐ Hundreds of words per day
- ☐ Clearly more than ten, but less than a hundred
- ☐ Around ten words per day
- ☐ A few words per day
- ☐ I do not write with my mobile device

4. What text entry method do you typically use when you write with your mobile device and how long have you been using that method? Answer with a number indicating how long you have used the method.

0 = My current main text entry method on mobile devices

1 = Still using occasionally

2 = Have used regularly, but not anymore

3 = A little

4 = Just tried

5 = Have not used at all

☐ Multitap

☐ Predictive text entry on a telephone keypad (T9)

☐ QWERTY keyboard on a physical keyboard

☐ Virtual keyboard, QWERTY keyboard on a touchscreen

☐ Swype

☐ Other, which method? _____

5. Do you feel that producing text using smartphone is hard and if so, why?

☐ Screen is too small

☐ Keys are too small

☐ Writing is too slow

☐ Avoiding typing errors is hard

☐ Other reason: _____

Experimenter fills:

Date & time:

participant:

Notes:

tilting-based text entry, kyselylomake

Taustatiedot

Ikä: _____ Sukupuoli: _____ Kätisyys: _____

Koulutus: ☐ Peruskoulu
 ☐ Ammattikoulu
 ☐ Lukio
 ☐ Alempi korkeakoulututkinto
 ☐ Ylempi korkeakoulututkinto

Englannin kielen taito:

☐ Heikko
☐ Välttävä
☐ Hyvä
☐ Erinomainen
☐ Äidinkieli

Käyttökokemus

1. Miten kuvailisit itseäsi tietotekniikan käyttäjänä?

- ☐ Kokenut käyttäjä, olen käyttänyt useita eri laitteita ja tulen yleensä toimeen niiden kanssa.
☐ Pärjään laitteiden kanssa, mutta minulla ei ole paljon kokemusta erilaisista laitteista.
☐ Olen usein tarvinnut apua ongelmissa, joita laitteet ja sovellukset tuottavat minulle.

2. Mikä on ensisijainen mobiililaitteesi (puhelin tai taulutietokone), anna valmistaja, nimi, malli ja näytön koko:

3. Arvioi kuinka paljon kirjoitat mobiililaitteellasi tyypillisesti päivän aikana:

- ☐ Yli tuhat sanaa päivässä
☐ Satoja sanoja päivässä
☐ Selvästi enemmän kuin kymmenen, mutta vähemmän kuin sata
☐ Noin kymmenen sanaa päivässä
☐ Muutama sana päivässä
☐ En kirjoita mobiililaitteellani

4. Mitä tekstinsyöttömenetelmää tyypillisesti käytät kirjoittaessasi mobiililaitteellasi ja kuinka kauan olet käyttänyt kyseistä tekniikkaa? Vastaa antamalla numero joka vastaa käyttöäsi kunkin menetelmän kohdalla

0 = Pääasiallinen syöttötekniikka tällä hetkellä mobiililaitteella kirjoittaessani

1 = Käytän tätä edelleen toisinaan

2 = Olen käyttänyt säännöllisesti, mutta en käytä enää

3 = Olen käyttänyt menetelmää vähän

4 = Olen vain kokeillut menetelmää

5 = En ole koskaan käyttänyt

☐ Multitap:

☐ Ennustava tekstinsyöttömenetelmä (T9):

☐ QWERTY-keypad, fyysinen näppäimistö:

☐ Virtual keyboard, QWERTY-näppäimistö kosketusnäytöllä:

☐ Swype:

☐ Muu, mikä? _____

5. Koetko tekstin tuottamisen matkapuhelimella vaikeaksi ja jos niin miksi?

☐ Näyttö on liian pieni

☐ Näppäimet ovat liian pieniä

☐ Kirjoittaminen on liian hidasta

☐ Virheitä tulee liian herkästi

☐ Muu syy: _____

Tutkimuksen tekijä täyttää

Päivämäärä ja aika:

Testihenkilö:

Huomioitavaa:

NASA TLX – Tilting keyboards experiment

Participant: **Method:** **Session:** **Date:**

Mental Demand

How mentally demanding was the task?

Very Low										Very High									

Frustration

How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low										Very High									

Task success

How successful were you in accomplishing what you were asked to do?

Perfect										Failure									

Speed of use

How fast you feel you were using the text entry method in question?

Very fast										Very slow									

Ease of use

How easy it was to use the method in question?

Very easy										Very hard									

Learnability

How easy it was to know how to use the method in question?

Very easy										Very hard									

Informed consent form

By signing this form you verify that you have understood that:

1. Participation in this experiment is voluntary
2. The object of study is the text entry methods, not the participant.
3. You may discontinue your participation in the experiment at any time for any reason.
4. You may also withdraw your data from the experiment. Another participant will be recruited to take your place.
5. The data you contribute will be used for scientific purposes. This means that it may be published in scientific publications in a form that makes it impossible to identify you. Usually only aggregate metrics over the whole group of participants are published.
6. The data will be saved by the researchers for many years. If other researchers are granted access to it, the names of the participants will be removed.

I agree to participate in this experiment and confirm that the researcher has explained the experimental procedure and informed me of the duration of the experiment.

Name (print)

date

Signature

Tilt keyboard, Instructions

Introduction

Tilt keyboard is a keyboard for mobile phones. It utilizes tilting of the phone for text input.

When the Tilt keyboard starts, one of the four keyboard pages, the start page, is shown. The start page and the other pages are shown in the Figure 1 below. 11 keys are shown on each page. One key is always the backspace and another is the space. The function of the other keys depends on tilting.

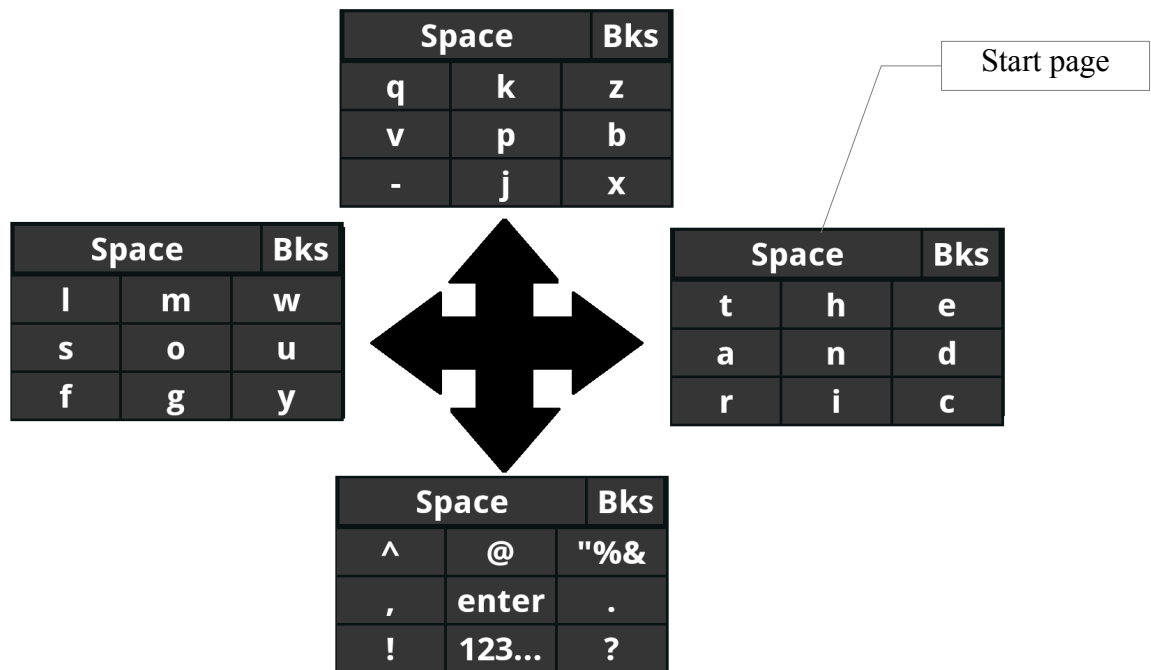


Figure 1. Tilt keyboard layout and tilting directions.

Using it

When you tilt the phone, a keyboard page with different set of characters/functions is shown. Tilting directions are left, right, up and down. Most common characters can be found from the right layout, second most common from the left layout and the least common ones from the up layout. The only button needed in this experiment from the down layout is enter. Character or function is selected by pressing a button with a finger.

When the Tilt keyboard starts the orientation/position of the phone is stored automatically. This is called the neutral position. The neutral position is a "safe zone" between the keyboard layouts where the layout does not change. Tilting is compared against the neutral position so that when the phone is tilted to one of four tilting directions from the neutral position, the layout changes accordingly. When typing, the phone should be tilted to the direction where the desired character is found and then returned to the neutral position for selecting the character/function.

The neutral position can be reset at any time by pressing the volume up button at the left side of the

phone while the phone is in the position that you want to be the new neutral position. You should reset the neutral position if you find yourself in a situation where the layout changes rapidly or not at all and you find it difficult to control. You should also reset the neutral position after pressing enter at the end of a phrase.

Tilter, Instructions

Introduction

Tilter is a keyboard for mobile phones operated by the combination of a button press and tilting of the phone. When the Tilter starts, one of the five keyboard pages, the start page, is shown. The start page has the space key and eight empty buttons. All the pages are shown in the Figure 1 below.

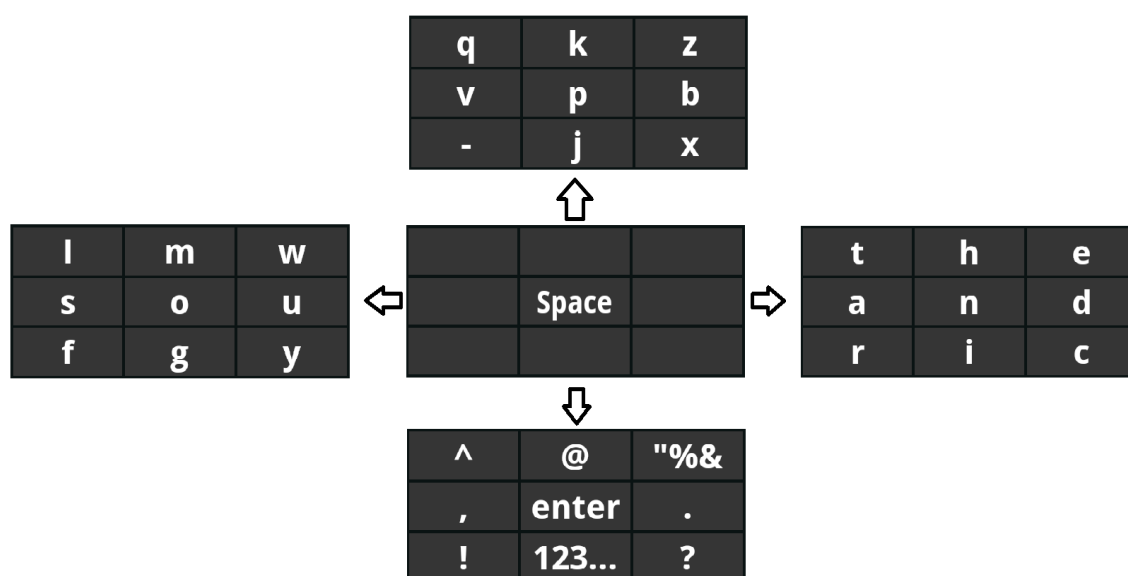


Figure 1. The Tilter keyboard layout and tilting directions.

Using it

You enter a character in four phases:.

1. Press the volume up button down (and keep it down)
2. Tilt the phone to select a keyboard page (and keep the volume button down).
3. Highlight the desired character on the page by tilting (see Figure 2 for example of a highlighted character).
4. release the volume button

t	h	e
a	n	d
r	i	c

Figure 2. Character 'e' has been highlighted.

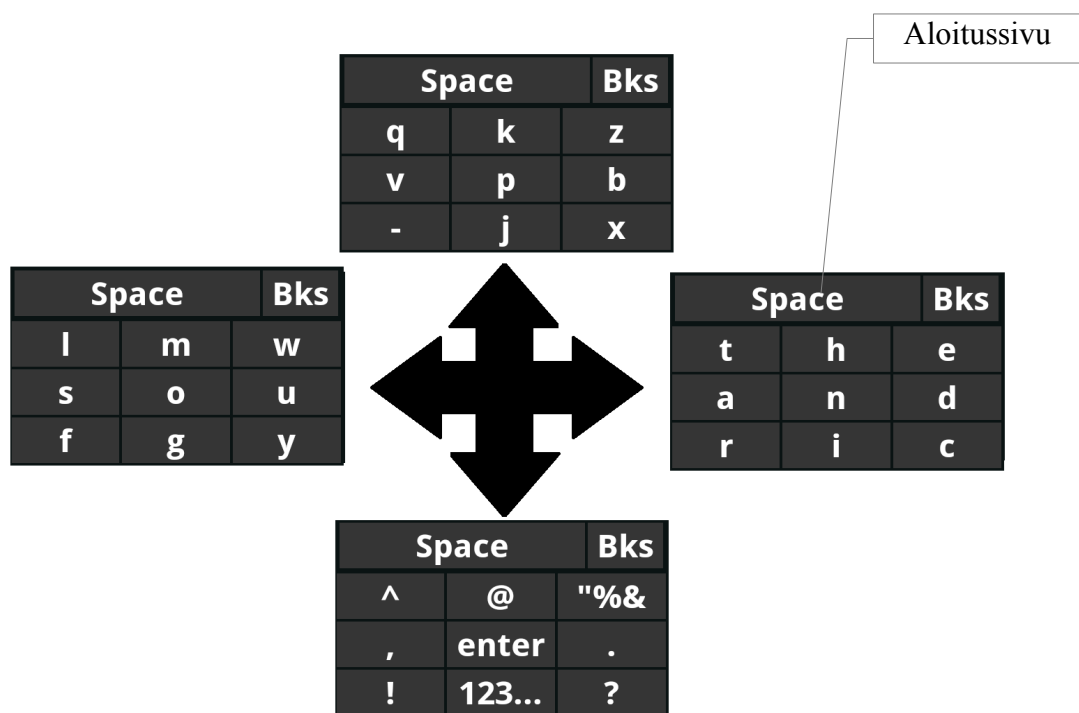
The volume up button is on the left side of the phone. Tilting directions are left, right, up and down. Most common characters can be found from the right page, second most common from the left page and the least common ones from the up page. The only key needed in this experiment from the down page is the enter. To input the space character you press and release the volume button. One of the nine buttons shown must be selected to continue after selecting a page. The mechanism to recover from an erroneous page selection is to enter a character and then press the volume down button to erase the entered character. When the volume up button is released, Tilter shows the start page.

Tilt keyboard, Ohjeistus

Esittely

Tilt keyboard on näppäimistö matkapuhelimelle. Näppäimistö hyödyntää puhelimen kallistelua tekstin syötössä.

Kun Tilt keyboard käynnistyy, aloitussivu, joka on yksi neljästä mahdollisesta näppäinsivusta, tulee näkyviin. Aloitussivu, sekä muut näppäinsivut, ovat nähtävissä kuvassa 1 alla. Kullakin sivulla on nähtävissä 11 eri näppäintä. Yksi näppäin on aina askelpalautin (backspace) ja toinen välilyönti (space). Muiden näppäinten toiminnallisuus vaihtelee kallistelun mukaan.



Kuva 1. Tilt keyboardin näppäinsivut ja kallistelusuunnat

Käyttö

Puhelinta kallistettaessa näppäinasettelu vaihtuu. Kallistussuunnat ovat vasen, oikea, eteen ja taakse. Käytetyimmät kirjaimet löytyvät oikeanpuoleisesta näppäinsivusta, seuraavaksi eniten käytetyimmät löytyvät vasemmanpuoleisesta näppäinsivusta ja vähiten käytetyimmät löytyvät eteenpäin suuntautuvasta näppäinsivusta. Ainoa näppäin, jota tässä kokeessa tarvitaan taaksepäin suuntautuvasta näppäinsivusta, on enter. Näppäin tai toiminto valitaan näppäimistöltä painamalla nappia sormella.

Kun Tilt keyboard käynnistyy, puhelimen asento tallennetaan automaattisesti. Tätä kutsutaan neutraaliasennoksi. Neutraaliasento on turva-alue eri näppäinsivujen välissä, jossa asettelu ei vaihdu. Kallistelua verrataan neutraaliasentoon siten, että kun puhelinta kallistetaan yhteen neljästä kallistelusuunnasta neutraaliasennosta aloitettaessa, näppäinasettelu vaihtuu. Kirjoitettaessa puhelinta kallistetaan siihen suuntaan, josta haluttu merkki löytyy, ja tämän jälkeen palataan neutraaliasentoon kirjaimen tai toiminnon valintaa varten.

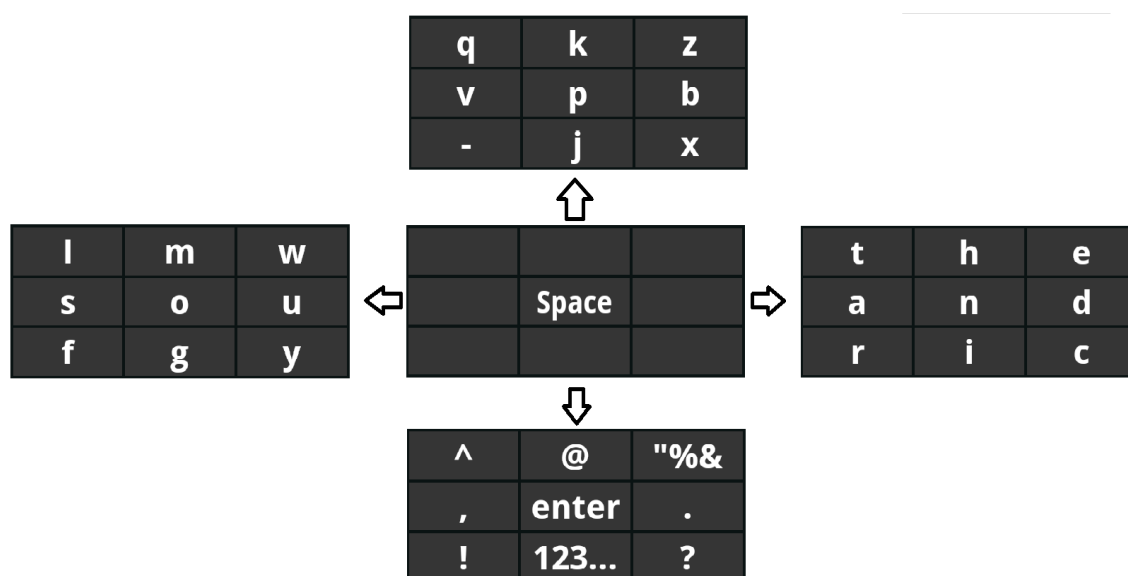
Neuraaliasennon voi uudelleen asettaa koska vain käytön aikana painamalla puhelimen volume up -näppäintä puhelimen vasemmalta puolelta puhelimen ollessa siinä asennossa, joka uudeksi neutraaliasennoksi halutaan.

Neutraaliasento tulisi uudelleen asettaa mikäli näppäinasettelu vaihtuu tiheään tai ei lainkaan. Näppäinasettelu tulisi uudelleen asettaa myös joka kerta kun testilauseen loppuksi on painettu enter -näppäintä ennen uuden lauseen kirjoittamisen aloittamista.

Tilter, Ohjeistus

Esittely

Tilter on näppäimistä matkapuhelimella, jota operoidaan yhden fyysisen näppäimen ja kallistelun yhdistelmällä. Kun Tilter käynnistyy, aloitussivu, joka on yksi viidestä näppäinsivusta, tulee näkyviin. Aloitussivulla on näkyvissä välilyönti (space) sekä kahdeksan tyhjää näppäintä. Kuvassa 1 alla, on näkyvissä kaikki näppäinsivut.



Kuva 1. Tilterin näppäinsivut ja kallistelusuunnat.

Käyttö

Merkki syötetään neljän eri vaihteen kautta:

1. Paina volume up -näppäin pohjaan ja pidä se pohjaan painettuna.
2. Kallista puhelinta haluttuun suuntaan valitaksesi näppäinsivu edelleen pitäen volume up -näppäintä pohjaan painettuna.
3. Korosta haluttu merkki kallistelun avulla (Kuvassa 2 on esimerkki, jossa kirjain on korostettu).
4. Vapauta volume up -näppäin merkin valintaa varten.

t	h	e
a	n	d
r	i	c

Kuva 2. 'e' -kirjain on korostettu valintaa varten.

Volume up -näppäin löytyy puhelimen vasemmalta puolelta. Puhelinta kallistettaessa näppäinasettelu vaihtuu. Kallistus suunnat ovat vasen, oikea, eteen ja taakse. Käytetyimmät kirjaimet löytyvät oikeanpuoleisesta näppäinsivusta, seuraavaksi eniten käytetyimmät löytyvät vasemmanpuoleisesta näppäinsivusta ja vähiten käytetyimmät löytyvät eteenpäin suuntautuvasta näppäinsivusta. Ainoa näppäin, jota tässä kokeessa tarvitaan taaksepäin suuntautuvasta näppäinsivusta, on enter. Välilyönti valitaan painamalla volume up pohjaan ja vapauttamalla se kallistelematta puhelinta.

Yksi yhdeksästä näkyvässä olevasta näppäimestä tulee valita käytön jatkamiseksi kun sivu on valittu. Mikäli valittu sivu on väärä, ja syötetty merkki ei haluttu, tulee syötetty merkki pyyhkiä pois. Tämä tapahtuu painamalla volume down -näppäintä. Tiltter näyttää aloitussivun joka kerta, kun volume up näppäin vapautetaan.